

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_nup model

Run on: June 21, 2004, 12:28:26 ; Search time 34.7285 Seconds
(without alignments)
11051.549 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532
Perfect score: 3615
Sequence: 1 atgagccatgagatgagtcg.....tggaaatcatgagtcctc 1995

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_nup.model -DEV=xldp
-Q=/cgn21/USPTO.spool.p/US10029345/runac_21062004_122816_4205/app_query.fasta_1.3278
-DB=pir_78 -QWMT=fastan -SUFFIX=rxp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10029345_QCGN_1_1_78@runac_21062004_122816_4205 -NCPU=6 -ICPU=3
-NO_MMAP -JARGOUDRY -NEG_SCORES=0 -WAIT -DSBLOC=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

pir_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	515	14.2	619	2	T15969
2	443	12.3	394	2	A56115
3	433	12.0	367	1	S24411
4	423	11.7	367	1	S52285
5	420	11.6	367	1	S29090
6	418	11.6	314	1	A57126
7	409	11.3	314	1	B57126
8	393.5	10.9	393	2	A56947
9	360.5	10.0	384	1	T13880
10	345	9.5	303	2	T16405
11	294.5	8.1	365	2	T132494
12	239.5	6.6	272	2	T18915
13	233.5	6.5	226	2	T13180
14	225.5	6.2	186	2	T16056

15.	223.5	6.2	330	2	T139698	protein tyrosine p
16	218.5	6.0	771	2	T47666	phosphatase-like p
17	208.5	5.8	220	2	T37885	low-molecular-mass
18	205.5	5.7	866	2	P88481	protein C16A3.1 [1
19	205	5.7	223	2	T139165	protein tyrosine p
20	204	5.6	278	2	T39517	dual-specificity p
21	202	5.6	344	1	T03074	dual specificity p
22	202	5.6	364	1	S31304	protein-tyrosine-p
23	196	5.4	185	1	A47196	dual specificity p
24	195	5.4	223	2	G84458	probable protein p
25	193	5.3	205	2	T149364	protein tyrosine p
26	189	5.2	489	1	S58725	dual specificity p
27	185	5.1	807	1	S44538	probable protein-c
28	183.5	5.1	276	2	T48906	protein-tyrosine-p
29	178	4.9	204	2	T17802	hypothetical prote
30	177	4.9	580	2	T18439	hypothetical prote
31	172	4.8	600	2	T18446	hypothetical prote
32	170.5	4.7	292	2	S41012	hypothetical prote
33	160	4.4	1132	2	T49403	related to protein
34	157	4.3	209	1	S48459	probable dual spec
35	154.5	4.3	5262	2	T03454	ALR protein - huma
36	154	4.3	1189	2	S56852	hypothetical prote
37	151.5	4.2	597	1	S43743	probable dual spec
38	151	4.2	1141	2	A44093	CGMP-inhibited CAM
39	150.5	4.2	169	2	T30684	probable dual spec
40	148.5	4.1	574	2	T43556	Wiskott-Aldrich sy
41	147	4.1	1168	2	S49915	extensin-like prot
42	146.5	4.1	272	2	T19418	hypothetical prote
43	145.5	4.0	171	1	T36845	dual specificity p
44	145.5	4.0	171	1	T28522	probable dual spec
45	145.5	4.0	171	2	B72161	J1L protein - vari

ALIGNMENTS

RESULT 1

T15969 hypothetical protein F08B1.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C/Accession: T15969

R/Chisoe, S.

submitted to the EMBL Data Library, July 1995

A/Description: The sequence of C. elegans cosmid F08B1.

A/Reference number: Z18439

A/Accession: T15969

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-619 <CHI>

A/Cross-references: EMBL:U03178; NID:G726421; PID:G726422; PIDN:AC46719.1; CESP:F08B1.1

A/Experimental source: strain Bristol N2

A/Gene: CESP:F08B1.1

A/Introns: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3

Alignment Scores:

Pred. No.: 7.36e-24 Length: 619
Score: 515.00 Matches: 165
Percent Similarity: 45.32% Conservative: 72
Best Local Similarity: 31.55% Mismatches: 173
Query Match: 14.25% Indels: 113
DB: 2 Gaps: 19

US-10-029-345A-108_COPY_538_2532 (1-1995) x T15969 (1-619)

QY 367 GGTGGTTTCTGAGTCTCTCGTTTCCCTGCGCTCTGTGAGAGAAATCC----- 420
DB 99 GTCGCTGCTGAGTCTCTCGTTTCCCTGCGCTCTGTGAGAGAAATCC----- 420
QY 421 ACTTGAATCCCTGAGTCTCTCGTTTCCCTGCGCTCTGTGAGAGAAATCC----- 420
DB 119 ThirArgLeuProGlnSerLeuSerGlnProCysLeuSerGlnProThrArgLeu---Gly 137

QY	475	CCAGCCGGAATCTTCCCAATCTTTATCTTGGCTGCCAGAGATGCTCAACAGAGAG	53
Db	138	lIethnleuilethProasnlleTyrlenglySerGlnlIeAspSerleuAspGlnThr	15
QY	535	CTGATACAGACAGAAATGGATGGTATATGTTAAATGACAGTAACTGTGCCAAAGCT	59
Db	158	MetleuAspAlaIeuaAspIleSerValIalleuAsnleuSerMetThCysProIySer	177
QY	595	GACTTATCCCGAG---TCTCATTTCCCGGTGGTGGCTGTGAATGACAGCTTTTGGAG	651
Db	178	ValCysIleIysGlnAspIlyAsnIheMetArgIleProValIaAspSerTyGlnGlu	197
QY	652	AAATTTTGGCCGTGGTGGACAAATCACTAATGATCTTATGGAAGAAAGAAAGCTCCAA	711
Db	198	LysIeuSerProTyPheProMetAlaTyrluIheuGlnIyIysCysArgAlaGlnGly	217
QY	712	GGATGTCTTATAGTCATGTTTATAGTGGATCTCCGCTCCGCCACCATGCTATGCGC	771
Db	218	LysIyScysIleuIleHisCysIeuAlaGlyIleSerIyIySerProThrIeuAlaIleSer	237
QY	772	TACATCATGAAGAGAGATGACATGCTCTTATGATGAACCTTACAGATTTTGGAAGAAAA	831
Db	238	TyrlIeMetArgTyImetIySmeGlySerAspAspAlaTyArgTyValIySglnAsy	257
QY	832	AGACCTACTATATCTCCAACTTCAATTTTCTGGGCCAACTCTCGATGTAGAGAAGAG	891
Db	258	ArgProSerIleSerProAsnPheAsnPheMetGlyGlnIeuIeuGlnIyTyGlnuIenVal	277
QY	892	ATTAAAGAACAGACTGGAGCATGAGGCCAAAGACAAATCAAGTCTGTGCACTCGAG	951
Db	278	Leu-----IleuAspIheValIeuAspTyArgAsn	287
QY	952	AAGCAATCAACCTGTCCCTGCTGTCTCAGAGAGGTGACAGAAAGACAGACGCCCTTC	1011
Db	288	GlnIaSerArgProHisArgHisMetAspTyTyTyGlyProSerAspIeuIySerPro---	306
QY	1012	AGTCCACCCCTGTCCCACTGTGTACTCTCAGAGGACAGACAGAAAGCCCTGATCTCC	1071
Db	307	---ProIyValProIySerAlaSerIaSerAsnCysValIheProGlySerThrlHisAsp	325
QY	1072	GCCAGCGTGGCCAGCGCTGCCAGCCGTGCAGCCGTGCTTTAGAGAGACAGCCGGTGTA	1131
Db	326	GlnSerSerProSerSerProSerValSerGlnGlySerAlaIaSerGlnPro-----	343
QY	1132	CAGGCGCTCAGTGGGTGACCTGTCCGAGACAGCTGGAAAGACAGCAATTAAGCTCAAG	1191
Db	343	-----	343
QY	1192	CGTCTCTCTCTGGATATCAATCAATTCATTTATATGAGCGACAGATGCGCATCTTA	1251
Db	344	-----GlnThrSerSerIaIaIaSer---	351
QY	1252	CATGGCTTCTCTCATTCAGAAAGATCCTTTGGAATTAACAACCTTCCACTTACTTGAT	1311
Db	352	-----SerSerSerThrlAlaSerIaIaProProSerMetProSerThrlSerGlnGln	368
QY	1312	GGAGCAACCAAGCTATGCAAGTCTCCCTGTTCAGAACTATGAGAGAGACAGTCCCGAA	1371
Db	369	GlyThrSer-----SerGlyThrAlaSnValIaSnGlyIySArgAsnMet	383
QY	1372	ACCAAT-----CTGATTAAGAGGAGAACCC---AGCATCCCAAGAAAGCTGACG	1416
Db	384	ThrlMetAspIeuGlyIeuProHisArgProIyValIaIeuGlyIeuProSerArgIleGly	403
QY	1417	ACGCGC-----AGCGCTTTCAGACAGC---CAGAGCAAGAGATTTGATTCGGTCAAA	1464
Db	404	ThrlSerValAlaGlnIeuProSerProSerThrlGlnIeuSerArgIeuSerPheAsnGln	423
QY	1465	ACCAAGAGCAAGTGGACAGCCGCCAGAGGTCCCTTTATCTTCACTGCATGCAAGTGGAGAG	1524
Db	423	YProGlnAlaIleIaIaProSerThr---ProIleIeu-----	434
QY	1525	GTCGAGGACAAATTACACACAGCTTCTTTTGGGCTTTCCACAGCCAGACAGACCTTC	1584

```

Db          435  -----AsnPhetThrAsnProCysPheAsnSerProIleIleProValAlaSerSerSe 452
QY          1585  ACGAAGTCGTGCGGGCGCTGGGCTTTAAGGGCTGGACACTGGATATCTTGAGCCCCCAAGACC 1644
           |||  :::::
Db          452  TargIuValIle----- 456
QY          1645  TTACCCCTTCGCTGACAGCAGCTGTATTTGGCAGAGTCTCTCACTTCTACTCT 1705
           |||  |||  |||  |||
Db          457  -LeuThrIeuProThrProIaIaIa-----SerSerSerSerIleIle 471
QY          1705  GCCTCAGCA-----TCTACGAGAGCGAGTGCACAGT 1734
           ::|||
Db          471  rSerIInProSerPheAspPheSerSerPheGluSerSerSerSerSerIleVala 491
QY          1735  TACTTGCCCTTCAGCTGAGCAGCTGCCACTTGGCGAAGCCAAAGTCTATTTCTGTGGCC 1794
           |||  |||  :::::  |||  ::|||
Db          491  IgluAsnProPhePheAlaSerIInGluValProIaGlySerSerSerIleSerThrPr 511
QY          1795  AGCGGGCGAAGGCCAAGTACAGACGACTGCCGCGGAGCTGCATGAAGAGAGGCC 1854
           |||  |||  |||  :::::  |||
Db          511  oSerGly---SerGlnSerThrProAlaSerAlaSerSerSerAlaIaIaSerIInGlySar 530
QY          1855  TTGTAAA 1861
           ::|||
Db          530  gMetLys 532

```

RESULT 2

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 4 - human
 N:Alternate names: dual specificity phosphatase HVH2
 C:Species: Homo sapiens (man)
 C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 24-Apr-1998
 C:Accession: A56115
 R:Guan, K.-u., Butch, E.
 J. Biol. Chem. 270, 7197-7203, 1995
 A:Title: Isolation and characterization of a novel dual specific phosphatase, HVH2, which
 A:Reference number: A56115, PMID:95221370; PMID:7535768
 A:Accession: A56115
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-394 <GMA>
 A:Cross-references: GB:U21108
 C:GeneIDs:
 A:Gene: GDB:DUSP4; HVH2; MKP-2
 A:Cross-references: GDB:433893
 A:Map position: 8p21-8p11.2
 C:Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity
 C:Keyword: phosphoprotein; phosphoric monoester hydrolase
 F:303-334/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
 F:380/Active site: Cys (phosphoserine intermediate) #status predicted
 F:386/Binding site: substrate phosphate (Arg)#status predicted

Alignment Scores:

File No.:	length:	394
Score:	Matches:	120
Percent Similarity:	Conservative:	74
Best Local Similarity:	Mismatches:	142
Query Match:	Indels:	46
DB:	Gaps:	14
		2

OS-10-029-345A-108_COPY_538_2532 (1-1995) X A56115 (1-394)

[illegible]

RESULT 3

Dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - mouse

Db	77	AlaIyselYserValSerLeuGlnGlnIleLeuProAlaGluIuGluValArgAlaArg	96
Oy	229	GTTGACATGATTGGCATCAGAAAGTGTGAATTGATTCAGATCAAAAGCTCCCAAGTGTGCC	288
Db	97	LeuArgSerGlyLeuThrYrSerAlaValIleValIlyTrsPrLuarXerProArgAlaGlu	116
Oy	289	TCTCTCTTCACACTGTTTTTCTCACCTGTACTTCTGGSTAAATGGAGAAGACTTC---	345
Db	117	SerLeuArgGluAspSerThrValSerLeuValValGlnIleLeuArgAAspAlaGlu	136
Oy	346	---AACCTGTTCACCGTTTGACAGGAGGGTGTGCTAGTCTCTGCATGTTCCCTGCAG	402
Db	137	ArgThrAspIleCySLeuLeuLySgilylTyGlnArgPheSerSerGutryrProGlu	156
Oy	403	CTCTGTGAGAAAATCCACTTA-----GTCCCTACTGCATTTCT	444
Db	157	PheCySerIyrThrlYsAlaLeuAlaIleProProProValProProSerAlaThr	176
Oy	445	CAGCCT-----TGC---TTACCTGTGCCAAAT---GGGCCA	477
Db	177	GluProLeuAspLeuGlyCySerSerCySgilylThrProLeuHisAspGlnGlyGlyPro	196
Oy	478	ACCAGATTCTTCCCACTTTATCTTGAGGTGGCCAGAGAGATGCTCCAAGAGAGCTGAC	537
Db	197	ValGluIleLeuProPheLeuTyrleneGlySerAlaTryhisAlaAlaArgAAspMet	216
Oy	538	ATACAGACAGATGGATGTGTTATGTGTATAATGCCACTATACCTGTCCAAAGCTGAC	597
Db	217	LeuAspAlaLeuGlyIleThrAlaLeuLeuAsnValSerSerAspCyProAsn---His	235
Oy	598	TTTATCCCCAGTCTCATTTCTGCGTGGCTGTGATACACACTTTGTGAAAAATT	657
Db	236	PheGluIylHisTyrlGlnTyrlYsCyAlleProValGluAspHisIylYsAlaAspIle	255
Oy	658	TTGCGCGGTGGACAATCAGTAGATTATTCATGGAACAAAGCAAGCTCCATGGATGTG	717
Db	256	SerSerThrPheMetGcluaIalleGutryrIleAspAlaValYsAspCyAspGlyAsp	275
Oy	718	GTTCTAGTGCACGTGTTTAGCTGGAGATCTCCCGCTCCGCCACATCGCTATCGCTACATC	777
Db	276	ValLeuValHisCySglnAlaGlyIleSerArgSerAlaThrIleCySleuAlaTyrlen	295
Oy	778	ATGAGAGAGATGCATGTCTTAGATGAAGCTTACAGATTGTGTAAAGAAAAGACT	837
Db	296	MetCetylYslyAspValArgLeuGlnGluAlaPheGluPheValYsGlnArgArgSer	315
Oy	838	ACTATATCTCCAACTTCATTTCTGGGCCAATCTCTGACATATGAGAAGAAGATTAG	897
Db	316	IleIleSerProAsnPheSerPheMetGlyGlnLeuLeuGlnPheGlnSerGlnValLeu	335
Oy	898	AACACAGACTGGAGATCAGGGCCAAAGAGAAACTCAAGCTGCTGCACCTGAGAGAGCCA	957
Db	336	AlaThrSerCySAlaAlaGluAlaIleSer-----Pro	346
Oy	958	AATGACACTGTCTCCGTGTCTCAGAGAGGTGACAGAAAAAGAGAGCCCTCAGTCCA	1017
Db	347	SerGlyProLeu-----ArgGluArgGlyIyrThrProAlaThrPro	360
Oy	1018	CCCCTGTCCGACCTGCTACTCCTCAGAGGACAGACAGAAAGCCCGTGCATCCCGCCAG	1077
Db	361	-----ThrSerGlnPheValPheSerPheProVal---SerValGly	373
Oy	1078	GTGCCAGAGCGGCCAGAGGTGACCGCTGCTGTTAAGACAGCCCGCTGTACAGAGCG	1137
Db	374	ValHisSerAlaProSerSerLeuProTyrlen-----HisSerProIleThrThrSer	391
Oy	1138	CTCAGT 1143	
Db	392	ProSer 393	

<p>N:Alternate names: 3CH134 protein; protein-tyrosine-phosphatase <i>erp</i>, nonreceptor type 10 C:Date: 19-Feb-1994 #sequence revision 02-May-1994 #ext_change 11-Jun-1999 C:Accession: A54681; S24411 R: Noguchi, T.; Metz, R.; Chen, L.; Matei, M.G.; Carrasco, D.; Bravo, R. Mol. Cell. Biol. 13, 5195-5205, 1993 A:Title: Structure, mapping, and expression of <i>erp</i>, a growth factor-inducible gene encoding A:Reference number: A54681; MUID:93360956; PMID:8355678 A:Accession: A54681 A:Molecule type: DNA A:Residues: 1-367 <NOG> A:Cross-references: GB:S64851; NID:g409976; PIDN:AAB27882.1; PID:g409977 R:Charles, C.H.; Adler, A.S.; Lau, L.F. Oncogene 7, 187-190, 1992 A:Title: cDNA sequence of a growth factor-inducible immediate early gene and characterized A:Reference number: S24411; MUID:92158357; PMID:1741163 A:Accession: S24411 A:Molecule type: mRNA A:Residues: 1-367 <CHA> A:Cross-references: EMBL:X61940; NID:g49735; PIDN:CAA43944.1; PID:g49736 C:Genetics: A:Gene: <i>erp</i> A:Introns: 123/1; 172/1; 245/1 C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity C:Keywords: immediate-early protein; phosphoprotein; phosphoric monoester hydrolase F:181-312/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL> F:258/Active site: Cys (phosphocysteine intermediate) #status predicted F:264/Binding site: substrate phosphate (Arg) #status predicted</p>	
<p>Alignment Scores: Pred. No.: 7,07e-19 Length: 367 Score: 433.00 Matches: 115 Percent Similarity: 46.43% Conservative: 54 Best Local Similarity: 31.59% Mismatches: 151 Query Match: 11.98% Indels: 44 DB: 1 Gaps: 8</p>	
<p>US-10-029-345A-108_COPY_538_2532 (1-1995) x S24411 (1-367)</p>	
QY	43 TTGGGCGCTGCTGCGAAAGTGGAAAGCGAAAGTCTGCTAAATTGATGACCGGCATT 102
DB	13 LENAAGLALALEULEAARGLGLGLYALALAGLNCYLEUULEUENPYSARGSERPHE 32
QY	103 GTGGAATACATATCATCCCACTTTTGAAGCCATTATATCACTGCTCCACGTTATG 162
DB	33 PHEALAPHEANALAGLYHISLEALAGLYSERVALASNAVARGPHESETHIRLEVAL 52
QY	163 AAGCAAGGTTCCACACGACGAAAGTGTATATACAGAGCTCATCCGACATTCGCGAA 222
DB	53 ARGAGARGALALALYGLYALAMETGLYLEUGLHISILEVALPROBENALAGLULEARG 72
QY	223 CATAGAGTTGACATTGATGACAGTGCAGAGTGGTGTATTACAGTCAAGCTCCCAAGAT 282
DB	73 GLYARGLEULEALAGLYALATYRHSIALAVALEULEUENPGLIARGSERALASER 92
QY	283 GTTGCGCTCTCTCTTCAGACCTGTTTCTCACTGTACTTCTGGGTAAACTG-----GAG 336
DB	93 LEUARGLYALALAYARGAPRGLYTHRLEUALALEUALALAGLYALALEUCYARGGLU 112
QY	337 AAGAGCTCAACTCTGTTTCACTGCTTGCAGAGTGGTGGTGGTGGTGGTGGTGGTGGTTC 396
DB	113 ALAARGSETHRGLNVALPHEPHELEUNGNGLYTYRGLIUALAPHESERALASERCYS 132
QY	397 CTTGGCGCTGTGAGAAATTCACACTTAGTCCCACTGCAATTTCTCAGCCT----- 450
DB	133 PROGLULEUCYSERLYGINSERTHR-----PROTHGLYLEUENPGLIARGSERALASER 150
QY	451 -----TGGTAACTGGTCCCAACATT----- 471
DB	151 THRSEVALPROAPSERALAGLUSERGLYCSERSECYSERCTHRPROLEUTYRASP 170
QY	472 -----GGGCCAACCCGAATTCTTCCCAATCTTATCTTGGCTGCGACGAGATGTCTC 525

Db 171 GlnGlyClyProValGlnIleLeuSerPheLeuTyrLeuGlySerAlaTyrThrIleAspSer 190

QY 526 AACAGAGAGCTGATACAGCAAGATGGAGTGGTTATGTGTTAAATGCCACCTATACCTG 585

Db 191 ArgLysAspMetLeuAspAlaLeuGlyIleThrAlaLeuIleAsnValSerIleAsnGly 210

QY 586 CCAACACCTGACTTATATCCCGAGTCTCATTTTCCGTGCGTGGACCTGGAAATGACACTTT 645

Db 211 ProAsn---HisPheGluGlyHisTyrGlnIlyrTySerIleProValGluAspAsnHis 229

QY 646 TGGAGAAATTTTCCCGTGGTGGACAAATCAGTATTCATTTGAGAAAGCAAAAGCC 705

Db 230 LysAlaAspIleSerSerTyrPheAsnGluAlaIleAspPheIleAspSerIleLysAsp 249

QY 706 TCCAAATGAGATGTGTTCTAGTCACCTGTTAGCTGGGATCTCCCGCTCCGCCACCATCGCT 765

Db 250 AlaGlyGlyArgValPheValHisCysGlnIleGlyIleSerTyrSerIleThrIleCys 269

QY 766 ATCGGCTCATCATGAGAGAGATGACATGTCTTTGATGAAGCTTACAGATTTGTGGAA 825

Db 270 LeuAlaTyrLeuMetArgThrAsnArgValIlyLeuAspGluAlaPheGluPheValIly 289

QY 826 GAAAAAAGACCTATATCTCCAAATCTCAATTTTCTGGGCAATCTCTGGACTATGAG 885

Db 290 GlnArgLysSerIleIleSerProAsnPheSerPheMetGlyGlnLeuGlnIlePheGlu 309

QY 886 AAGAAGATTAAAGAACAGACTGAGCATGAGGCGCAAAAGCAAACTCAACTCTGTGAC 945

Db 310 SerGlnVal-----LeuAlaProHisCysSer----- 318

QY 946 CTGAGAAAGCCAAATGAACCTGCTCCCTGCTGTCTAGAGGCTGAGCAAAAAAGCAAGAC 1005

Db 319 ---AlaGluAlaGlySerProAlaMetAlaValLeuAspArgGlyThrSerThrThr 337

QY 1006 CCCCTGATGCCACCTGTGCTCGCACTGTGCTACTAGAGGACAGACAAAGCCCGTG 1065

Db 338 ValPheAsnPheProValSer-----IleProVal 347

QY 1066 CATCCGCCACAGC 1077

Db 348 HisProThrAsn 351

RESULT 4

S52265
dual specificity phosphatase (EC 3.1.3.-) 1 - rat
N/Alternate names: protein-tyrosine-phosphatase Ctl100; protein-tyrosine-phosphatase, not C/Species: Rattus norvegicus (Norway rat)
C/Date: 07-MAY-1995 #sequence_revision 21-Jul-1995 #text_change 22-Jun-1999
C/Accession: S52265
R/Mda, M.; Schlegel, W.; Arkinstall, S.
submitted to the EMBL Data Library, January 1995
A/Description: Pathways regulating Ctl100 gene expression in pituitary cells.
A/Reference number: S52265
A/Accession: S52265
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-367 <MUD>
A/Cross-references: EMBL:X84004, NID:G642264, PID:CAA58828.1, PID:G642265
C/Keywords: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity P;181-312/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl> P;185/Active site: Cys (phosphocysteine intermediate) #status Predicted
P;126/Binding site: substrate phosphore (Arl)

Alignment Scores:	
Pred. No.:	2.88e-18
Score:	42.00
Percent Similarity:	46.05%
Best Local Similarity:	31.34%
Query Match:	11.70%
DB:	2
Length:	367
Matches:	115
Conservative:	54
Mismatches:	148
Indels:	9
Gaps:	0

US-10-029-345A-108_COPY_538_2532 (1-1995) X S52265 (1-367)

QY 43 TTGGTGGCTCTCTGGAAAGTGAACGGAAAGAAAGTGGTGTATTAATGATGACCGGCAATT 1
 Db 13 LeuAlaGlyAlaLeuLeuAlaGlyAlaGlyAlaGlyLeuLeuLeuLeuLeuSerPhe 3
 QY 103 GTGAATACAAATATCATCCCAATTGTGGAGCCATTATATCAATCTGCCAAAGTTAT 1
 Db 33 PheAlaPheMetAlaGlyHisIleValGlySerValAsnValAlaPheSerThrIleVal 5
 QY 163 AAGCGAAGCTTGCAA-----CAGGACAAAGCTTTAATTAACAGAGCTATC 2
 Db 53 ArgAlaGlyAlaIaIysGlyAlaMetGlyLeuGlnHisIleValProAsnThrGluLeu--- 7
 QY 208 CAGCATTCACGGAAACAATAGCTTGACATTGATTCGACGTGACAGAGCTTGATTCAGAT 2
 Db 72 -----ArgGlyArgLeuLeuAlaGlyAlaIaIyTrnHisAlaValValLeuLeuAsp 8
 QY 268 CAAGAGCTCCCAAGATGTGGCTCTCTCTCTTTCAGATGTGTTCACGTGTTCTACGTACTTGGGT 3
 Db 88 GluTrgSerIaAlaLeuAspGlyAlaIaIysArgAspGlyTrnLeuAlaLeuAlaAlaGly 10
 QY 328 AAATCG-----GAGAAAGAGCTTCAACTGTGTTCACCTGTTCAGAGGTGGCTTGCTGAG 38
 Db 108 AlaLeuCysArgGlnAlaArgSerThrGlnAlaPhePheLeuGlnGlyIaIyTrnGluAla 12
 QY 382 TTCTCTCGTGTGTTCCCTGGCTCTGTGAAGAAA----- 41
 Db 128 PheSerIaSerCysPheGlnLeuCysSerIaIysGlnSerThrProMetGlyLeuSerLeu 14
 QY 418 -----TCCACTGTATGCTACCTGACCTGATTTCTCAGGCTTGCTTACCTGTGGCAACT 47
 Db 148 ProLeuSerThrSerValProAspSerIaGlnSerIaCysSerSerCysSerThrPro 16
 QY 472 -----GGGCGCAACCGAATTTCTTCCCAATCTTATCTTGAGTGGTCCAGAGCA 51
 Db 168 LeuIaIysArgGlnGlyIaIysProValGlnIleLeuSerPheLeuIaIyTrnGlySerAlaIaIyTrn 18
 QY 517 GATGTCCTCAACAGAGAGTATGACAGAAATGGAGTTGGTATATGTTAAATGCCAGC 57
 Db 188 HisAlaSerAlaGlyIaIysPheMetLeuAspAlaLeuGlyIleThrAlaLeuLeuHisAlaSer 20
 QY 577 TATACCTGTGCCAAAGCTGACTTTATCCCGAAGCTTCATTTCTCTGCGTGGCTGTGAAAT 63
 Db 208 AlaAsnCysProAsn---HisPheGlnGlyHisIaIyTrnGlnIaIySerIleProValGlu 22
 QY 637 GACAGCTTTGTGAGAAATTTTGGCTGGTGGAGCAATGAGTATGATTTATTTAGAA 66
 Db 227 AspAsnHisIaIysAlaPheIleSerSerTrpPheAsnGlnAlaIleAspPheIleAspSer 24
 QY 697 GCAAAAGAGCTCAATGATGTGTTCTTACGTGCACCTGTTAGCTGGAGATCTCCGCTCGGC 75
 Db 247 IleIaIysPheAlaGlyAlaIaIysArgValPheValHisCysGlnAlaGlyIleSerArgSerAla 26
 QY 757 ACCATGCCATATGCCCTTACATCAAGAAGAGATGACATGTCTTTAGATGAAAGCTTACAGA 81
 Db 267 ThrIleCysLeuAlaIaIyTrnLeuMetAlaGlnHisAsnArgValIaIysLeuAspGlnAlaPheGlu 28
 QY 817 TTTGTGAACAACAAAAGACCTATCTATCTTCCAAATTCGAATTTCTGGGCAACTCTG 87
 Db 287 PheAlaIaIysGlnArgTrgSerIleIleSerProAsnPheSerPheMetGlyGlnLeuLeu 30
 QY 877 GACTATGAGAAAGATTTAAGAACACAGACTGAGACATCAAGGCGCAAAAGCAAACTCGAG 93
 Db 307 GlnPheGlnSerGlnVal-----LeuAlaProHisCysSer----- 318
 QY 937 CTGCTGCACTGAGAGCAAAATGAACTGTCTCTGCTCTCTTCCAGAGGTGACAGAA 99
 Db 319 -----AlaGlnAlaGlySerProAlaMetAlaValLeuAspArgGlyHisSer 334
 QY 997 AGGAGAGCGCCCTCATGTCACCTGTGCGCAGACTTGCTACTCAAGAGCACACAGAGAA 1056
 Db 335 ThrTrnThrValPheAsnPheProValSer-----

QY 1057 AGGCCGTCATCCGCCAGC 1077
|||||
Db 345 ILeProValHisProThrAsn 351
RESULT 5
S23090
dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - human
N/Alternate names: protein-tyrosine-phosphatase Ctl00; protein-tyrosine-phosphatase, not
C/Species: Homo sapiens (man)
C/Date: 25-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999
C/Accession: S23090; #EMBL:U00002; #FBI:U00002; #FBI:U00002; #FBI:U00002
R/Key: S.M.; Emslie, E.A.
Nature 359, 644-647, 1992
A/Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine
A/Reference number: S23090; PMID:93024952; PMID:1406996
A/Accession: S23090
A/Molecule type: mRNA
A/Residues: 1-367 <KEY>
R/Key: S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.
J. Biol. Chem. 269, 3596-3604, 1994
A/Title: Isolation and characterization of a human dual specificity protein-tyrosine pho
A/Reference number: A53052; PMID:94148864; PMID:8106404
A/Accession: A53052
A/Molecule type: DNA
A/Residues: 1-367 <KEY>
A/Experimental source: Leukocyte
A/Note: sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804,
C/Genetic:
A/Accession: GDB:DUSP1; PTPN10
A/Map position: 5q34-5q34
C/Suprafamily: dual specificity phosphoprotein phosphatase 1, vhl-type dual specificity
C/Keywords: heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induced P
F/181-312/Domain: vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
F/258/Active site: Cys (phosphocysteine intermediate) #status predicted
F/264/Binding site: substrate phosphate (Arg) #status predicted
Alignment Scores:
Pred. No: 4.38e-18 Length: 367
Score: 420.00 Matches: 113
Percent Similarity: 46.69% Conservative: 56
Best Local Similarity: 31.22% Mismatches: 159
Query Match: 11.62% Indels: 34
Gaps: 7
US-10-029-345a-108_copy_538_2532 (1-1995) x S23090 (1-367)
QY 43 TTGGTGCTCTGCGAAAGTGAACGAAAGTCTGTAATGATGATCCGCCATT 102
|||||
Db 13 LeuArgAlaLeuLeuGlyGluArgAlaAlaGlnCysLeuLeuLeuSerPhe 32
QY 103 GTGGAATACATACATCCACATTTTGGAGCCATTAATATCACTGCTCAAGCTTATG 162
|||||
Db 33 PheAlaPheHisAlaGlyHisIleAlaGlySerValHisAlaArgPheSerThrIleVal 52
QY 163 AAGCGAAGTTGGCAACGACGAAAGTTTAATTAACAGCTCATCCAGCTTACGCAAA 222
|||||
Db 53 ArgArgArgAlaValGlyAlaMetGlyLeuGlnHisIleValProAlaAlaGluLeuArg 72
QY 223 CATPAGGTTGACATGATGTCAGTCAGCAAGAGTTGATTAACATCAAGCTCCCAAGAT 282
|||||
Db 73 GlyArgLeuLeuAlaGlyAlaValHisAlaValAlaLeuLeuSerPheGluArgSerAlaAla 92
QY 283 GTTGCCCTCTCTCTTCAGACTGTTTTCAGCTGACTTCGAGTAACTG-----GAG 336
|||||
Db 93 LeuArgPheAlaValArgSerPheGlyThrLeuAlaLeuAlaAlaGlyAlaLeuGluValGlu 112
QY 337 AAGAGCTTCAACTGTTGACCTGCTGCAAGTGAGGTTGCTGAGTTCTCTGTTGTTTC 396
|||||
Db 113 AlaArgAlaAlaGlnValPhePheLeuGlyGlyGlyGluAlaPheSerAlaSerCys 132
QY 397 CTGGGCTCTGTGTAAGAAAA-----TCCAGCTCTA 426

Db 133 ProGluLeuCysSerIleGlnSerThrPrometGlyLeuSerIleuProLeuSerThrSer 152
|||||
QY 427 GTCCCTACCTGCAATTTCTACGCTTCTGCTTACCTGTTCCCAACTT----- 471
|||||
Db 153 ValProAspSerAlaGluSerGlyCysSerSerCysSerThrProLeuThrArgGlnGly 172
QY 472 GGGCCAAACCGAATTTCTTCCCAATCTTTATCTTGGTGGCCAGGAGATGCTTCAACAG 531
|||||
Db 173 GlyProValGlnIleLeuProPheLeuThrGlySerAlaThrHisAlaSerArgGly 192
QY 532 GAGCTGATACAGCAAGATGATGTTGTTATGTTTAAATGCGATATCTGTCGAAG 591
|||||
Db 193 ArgMetLeuAspAlaLeuGlyIleThrAlaLeuHisAlaValSerHisAlaGlyProAsn 212
QY 592 CTTGACTTTATCCCGAGTCTCATTTCTGCGTGGTGGTGGTGGTGGTGGTGGTGG 651
|||||
Db 213 ---HisPheGlnGlyHisIleThrGlySerIleProValGlnAspHisIleVal 231
QY 652 AAAATTTTCCCGTGGTGGCAAAATGATGATGATGATGATGATGATGATGATGATGAT 711
|||||
Db 232 AspIleSerSerThrPheAsnGlnAlaIleAspPheIleAspSerIleValAsnAlaGly 251
QY 712 GGATGTTTCTAGTGCATGTTTGGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 771
|||||
Db 252 GlyArgValPheValHisCysGlnAlaGlyIleSerArgSerThrIleCysLeuAla 271
QY 772 TACATCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 831
|||||
Db 272 TyrLeuMetArgThrAsnArgValIleValAspGlnAlaPheGluPheValIleGlnArg 291
QY 832 AGACCTATATATCTTCAAACTTCAATTTTCTGGGCACTCTGATGATGATGATGATGAT 891
|||||
Db 292 ArgSerIleLeuSerProAsnPheSerPheMetGlyGlnLeuLeuGlnPheGlnSerGln 311
QY 892 ATTAAGAACAGACTGAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 951
|||||
Db 312 Val-----LeuAlaProHisCysSer-----Ala 319
QY 952 AAGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1005
|||||
Db 320 GluAlaGlySerProAlaMetAlaValLeuAspArgGlyThrSerThrThrValPhe 339
QY 1006 -----CCCTCAGTCCACCTGCTGCGGACTGCTGATCACTCAAGCAGCAGCAAG 1059
|||||
Db 340 AsnProValSerIleProValHisSerThrAsnSerAlaLeuSerThrLeuGlnSer 359
QY 1060 CCCGTG 1065
|||||
Db 360 ProIle 361
RESULT 6
A57126
dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 2 - human
N/Alternate names: mitogen-induced nuclear tyrosine phosphatase; phosphatase of activator
C/Species: Homo sapiens (man)
C/Date: 03-Nov-1995 #sequence_revision 02-Jul-1996 #text_change 11-Jun-1999
C/Accession: A57126
R/Key: R.P.; Davies, P.; Moskaluk, C.A.; Kearns, M.; Kruttsch, H.; Siebenlist, U.; Kelly,
Science 259, 1763-1766, 1993
A/Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
A/Reference number: A57126; PMID:93206122; PMID:7681221
A/Accession: A57126
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-314 <ROH>
A/Cross-references: GB:U11329; NID:G559539; PIDN:AAA50779.1; PID:G292376
C/Genetic:
A/Accession: GDB:DUSP2
A/Cross-references: GDB:139200
A/Map position: 2q11-2q11
C/Function:
A/Description: catalyzes the hydrolysis of peptidyl-phosphoserine, -phosphothreonine, and

C:Superfamily: dual specificity phosphoprotein phosphatase 1, VH1-type dual specificity
 C:Keywords: nucleus, phosphoprotein, phosphoric monoester hydrolase
 F:140-150/Region: nuclear location signal
 F:180-311/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
 F:257/Active site: Cys (phosphocysteine intermedate) #status predicted
 F:263/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:

Seed, NO.:	5,75e-18	Length:	314
Score:	418.00	Matches:	109
Percent Similarity:	50.83%	Conservative:	45
Best Local Similarity:	35.97%	Mismatches:	105
Query Match:	11.56%	Indels:	44
DB:	1	Gaps:	0

02-10-025-343A-108_COPY_538_2532 (1-1995) X A57126 (1-314)

```

QY 70 AAAAAAGGCGCATATTTGATGCCGCCATTGTGGAAACATACATCCCACTTTG 128
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 25 GGuargthrleuenuenuspCyargPrpheuenuLapheCysargrghlsvlaArg 44
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 130 GAAGCCATTAAATTCACACTGCTCCAGGCTTATGAGGAAAGG----- 171
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 45 AlaAlaArgProValPProTrrpsnaLaleuenuArgArghlaArgGlyPProfoLa 64
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 172 -----TTGCAACAGGACAAAGCTGTAAATTCAGACCTCATCCAGCATTCa 216
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 65 AlaValleuenaLacyleuLeuPProAsphArgAlaLeuArgThrArgLeuValArgGlyL 84
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 217 GCGAAACATTAAGGTTCACATTGCAATGCAAGAGGTTTGATTCAGATCAAGAGTCG 276
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
    85 Leu-

```

[illegible][illegible]

176 leuvenuroltyrluenehenuolysercyserhisersecasphleuclindlyneudln 195
 544 CAGAAAGGATGGTGTATGTGTAAAGCCAGCATATCCGTCGCAAGCCTGACTTATC 603
 196 AlGcsglylEthAlaValleuasnValserAlaseCypro----- 210
 604 CCCAGCTTCATTTCTCTCGGCGCCCTGTAAATGACAGCTTT 645
 211 -----AsnHisPheGluGlyLeuPheArgTyrLysSerIlePheValGluAspAsnGln 228
 244 -----

646 YGAGAAAAATTTTCCCTGGTTTGCAAAATCGTAGATTTCATTGTGAAGAAGCAAAGCC 705
229 MetValGluIleSerAlaTrpPheGlnGluAlaIleGlyPheIleAspTrpValLysaaN 248
706 TCCAAATGAGATGTTGTAGTGCACATGTTTAAGCTGGAGTCTCCCGCTCCGCACACTCGT 765
249 SerGlyGlyArgValLeuValHisGySglnAlaGlyIleSerAgsSerAlaThrIleCys 268
766 ATCCGCTCAATCAAGAGAGAGATGCACATGTCCTTTAATGTGAAGCTTACAAGATTGTGA 825

Gy	826	GAATAAAGACCTACTATTCGCAACTCATTTTCTGGGCCAACTCCTGGACTATGAG	885
	:	: ::	:: :
Db	289	GlnHArgGlyValIleSerProaspPheSerPheMetGlyInuLeuGlnPheGlu	308
	:	: ::	:: :
Gy	886	AAGACGATT	894
	:	: ::	:: :
Db	309	ThrGlnVal	311

RESULT 7

dual specificity phosphatase (EC 3.1.3.-) 2 - mouse
 B57126
 N:Alternate names: mitogen-induced nuclear protein-tyrosine-phosphatase PNC-1
 C:Species: Mus musculus (house mouse)
 C>Date: 03-Nov-1995 #sequence revision 03-Nov-1995 #next_change 24-Apr-1998
 C:Accession: B57126
 R:Kohan, P.U.; Davis, P.; Moskaluk, C.A.; Kearns, M.; Krutzsch, H.; Sibenlist, U.; Kelly,
 Science 259, 1763-1766, 1993
 A:Title: PNC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
 A:Reference number: A57126; MUID:93206122; PMID:7681221
 A:Accession: B57126

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-314 <ROH>
A:Cross-references: GB:L1130
C:Superfamily: dual specificity phosphoprotein phosphatase 1, Vhl-type dual specificity phosphatase 1
G:Keywords: nucleus; phosphoprotein; phosphoric monoester hydrolase
F:180-311/Bdomain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
F:257/Active site: Cys (phosphocysteine intermediate) #status predicted
F:263/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:

RefSeq. NO.:	2,036-17	Length:	314
Score:	409.00	Matches:	107
Percent Similarity:	51.18%	Conservative:	45
Local Similarity:	36.03%	Mismatches:	113
Query Match:	11.31%	Indels:	32
Gap:	2	Gaps:	7

OS-10-029-345A-108_COPY_538_2532 (1-1995) X B57126 (1-314)

[illegible]

85 Leu-----AlaArgIaIaValIeuLeuAspIeuSer 95
277 CAAGATGTCCTCTCTCTCAGACGCTGTTTCTACTGTTGGTAACTGAG 336
||| ||| |||
96 AlseValThrGluLeuProAspPolYProAlaHisIeuLeuLeuAlaIaLeuGln 115
||| ||| |||
337 AAGAGCTT-----AACCTGGTCACTGTTGACGAGGGCTTCTCAGTTC 384
||| ||| |||
116 HisGluMetArgIaGlyIProThrThrValCysPheIeuArgIaGlyI PheIeuSerPhe 135

[illegible]

Df 176 IIELEUPFOYRILEUYYRLEUGLISerCYsaMhiIserrSerAapleugIngluIn 195
Df 544 CAGATGGGATTGGTGTATGTGTAATGCACGAATACTTCGCAAGCCGTACTATC 603
Df 196 AlaCygiLIeThnrlAlaValIeuAsnValSerAlaseryProasn--HisPheglu 214
Df 604 CCCGAGTCTCATTTCCCGTCGCCGCTGTGAATGACAGCTTTTGAGAAATTTTGCCG 663
Df 215 GILyEUphehIaryLyserSerileProvalGlunspangInMetValIgulleSerAla 234
Df 664 TGGTTGGACAATTCAGTATTTCATTGAGAAAGCAAAGCCCTCAATCGATGTTCTTA 723
Df 235 TrpHeugIngluAlalIeserPheIIeaspserVallysaansergilyglYArgvalIeu 254
Df 724 GTGCAGTGTTAGCTGGGATCTCCCGCTCCGCCACAATCGCTATCGCTCAATCAGTAAG 783
Df 255 ValHiecygslnaIaglyIlIeserArgerSerAlahrhIecysleuAlatyLeuIllegln 274
Df 784 AGATGGACATGCTCTTATGATGAAAGCTTACAGATTGTGAAAAGAAAAAGCCTACTATA 843
Df 275 serHlsargValArgyleuAspGlualAphesaspPhevallysGlnArgarglyValIlle 294
Df 844 TTCTCCAACCTTCATATTTCTGGGCCAATCCTTGAGCTATGAGAAGAGATT 894
Df 295 SerProaenPheSerPheMetcilyGlnIeuengInIeuGIuthrGlnVal 311

RESULT 8
AS6947
dual specificity phosphatase (EC 3.1.3.-) HNH2 - rat
N/Alternate names: mitogen-activated protein kinase phosphatase 2
C/Species: Rattus norvegicus (Norway rat)
C/Dates: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 17-Mar-1999
C/Acession: A56947
J/Misra-Press, A.; Rlm, C.S.; Yao, H.; Robertson, M.S.; Stork, P.J.S.
J. Biol. Chem. 270, 14587-14596, 1995
A>Title: A novel mitogen-activated protein kinase phosphatase. Structure, expression, and
A/Reference number: A56947; MUID:95301550; PMID:1782322
A/Acession: A56947
A>Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-393 <MIS>
C/Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
C/Keywords: phosphoprotein; phosphoric monoester hydrolase
F/202-333/Domains: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
F/279/Active site: Cys (phosphocysteine intermediate) #status predicted
F/285/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:
Pred. No.:
Score: 1,81e-16 Length: 393
Percent Similarity: 393.50 Matches: 115
Best Local Similarity: 46.02% Conservative: 64
Query Match: 29.56% Mismatch: 153
DB: 10.89% Indels: 57
Gaps: 13

US-10-029-345A-108_COPY_538_2532 (1-1995) x AS6947 (1-393)

Df 46 GTGGCTGTGCTGGAAGGCGAAGCGAAGAAAGTGTGCTGAATTGATGACCGGCAATTTGTG 105
Df 38 LeuIlglyLeuIeuserglyLy-----LysCyseuIeuIeuAapCybaIrgproPhleu 55
Df 106 GAATACATATCATCCCATTTTGGAGCCATTTAATATCACTGTCACAGCTTATGAAG 165
Df 56 AlahIsserAlaglyTylllearglyserValenValargCybaenthrilleValarg 75
Df 166 CGAAGTGGACAAGCAAGAAAGTGTTAATTAACAAGTCATTCAGCATTCAGCAAAAT 225
Df 76 Argarg--AlalygLyserValserIeuGlnIleu-----ProAlaglglu 92
Df 226 AAGGT-----GACTTGATTGCACTGCAAGAGTTGATGTTAGAT 267
Df 93 GluValArgProCybaAlaleuAlaserThrArgIeuserserThmetThrGlnPro 112

[illegible]

J Biol. Chem. 270, 1156-1160, 1995

A>Title: Multiple dual specificity protein tyrosine phosphatases are expressed and regulated by growth factors

A'Reference number: A55332; MUID:95138103; PMID:7836374

A'Accession: J38890

A'Molecule type: mRNA

A'Residues: 1-384 <RBS>

A'CROSS-references: EMBL:U16996; NID:G642012; PIDN:AAB06261.1; PTD:G642013

R:ribhshh.T.; Bottaro, D.P.; Michieli, P.; Kelley, C.A.; Aaronson, S.A.

J. Biol. Chem. 269, 29897-29902, 1994

A>Title: A novel dual specificity phosphatase induced by serum stimulation and heat shock

A'Reference number: A55313; MUID:95050849; PMID:7961985

A'Accession: A55313

A'Molecule type: mRNA

A'Residues: 1-87 GHV, 12-70, R, 72-104, F, 107-362, RCLPQGSSEALMORPNPAKTMESAPQSOI

A'CROSS-references: CB:U15932; NID:9606971; PIDN:AAA64693.1; PTD:9606972

C:Genetics:

A:Gene: GDB:DUSPS

A:CROSS-references: GDB:385447

A:Map position: 10q25-10q25

C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity phosphoprotein phosphatase 1; phosphoric monoester hydrolase

F:186-317/DNA: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>

F:263/Active site: Cys (phosphocysteine intermediate) #status predicted

F:269/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:

Pred. No.: 1

Score: 1.85e-14 Length: 384

Percent Similarity: 360.50 Matches: 107

Best Local Similarity: 43.67% Conservative: 62

Query Match: 9.97% Mismatches: 143

Indels: 75

Gaps: 10

US-10-029-345A-108 COPY 538 2532 (1-1995) x I38890 (1-384)

OY 52 CTGCTGGAAAGTGAACGGAATAAGTGCTGTATTTATGACCAGCCATTGTCGAATAC 111
||| :
Db 15 LeuargylsgluAlaAlaAlaArgCysValValLeuaspCysArgProTyrLeuAlaIaph 34
112 ATATCATCCCAATTTTGGAAAGCCATTATATCAACTGCTCAAGCTTATGAAAGCGAAG 171
35 AlAlaSerAsnValArgglyserLeuAsnValAsmLeuAsnSerValValLeuArgArg 54
OY 172 TGCACAACGACAAAGTGTAAATTACAGAGCTATCCAGATTCCAGCAACATAAGGTT 231
:::
Db 55 AlaArglyglValAlaVal-----SerAlaArgTyrValLeu 66
232 GACATGATGTCAGCTAG-----AAGGTT 255
Db 67 ProAspglnAlaAlaArgAlaArgLeuLeuGlnGlnGlyglValAlaAlaVal 86
OY 256 GTTAGTTTACGATCAAGCTGCCAAGATGTGCTGCTCTCTTCAGACGCTTTTCTACT 315
87 ValValAlaLeuAspGlnGlySerArgHisTrpGlnIlySleuArgGlnGlnSerAlaAlaArg 106
OY 316 GTACTTTCGGGTAACGTGAGAAGAGCTTCAACTCT-----GTTCACCTGCTTSCA 366
||| :
Db 107 ValValAlaLeuThrSerLeuAlaCysLeuProAlaGlyProArgValTyrPheLeuIlys 126
OY 367 GGTGGGTTTGTCTGACTTCTCTGTTGTTTCCCTGGCCCTCTGT-----408
127 GlyIlyTyrGlnThrPheTyrSerGlnIlyrProGlnCysCysValaIspValIysProile 146
OY 409 -----GAAGGAAATTCACCTTAgTCCCTAACCTGCACTTTCAGCCTTGC 453
147 SerGlnGlnIlyslieglsSerGlnArgAlaLeuIleSerGlnCysgIylys-----153
OY 454 TTACTCTGTGCAACATT-----GGGCCCAACCCCAATT 486
Db 164 ---ProvalValasnaValSerTyrArgProAlaTyrAspGlnIlyglProValGlnIle 182

QY	487	CTTCCCAATCTTATATCTTGCGCTCCGACGAGATGTCCTCAACAGAGCTGATACGACG	54
Db	183	LeuProhLeuTyrlLeuGlySerAlaTyrlrhIsAlaSerLysCysGlnPheLeuAlaSn	20
QY	547	AATGGAGTTGGTATATGTTAAATGCGACTTACCTTCACAAAGCTGACTTATATCCC	606
Db	203	LeuHisIleThrAlaLeuLeuSerValSerAlaGlyThrSerGluAla---CysMetThr	22
QY	607	GAGTCATCTTCTCGCGTGGCTGTGATGACGACTTGTGAGAAATTTTCCCGTGG	666
Db	222	HisLeuHisAlaTyrlsThrIleProValGlnubSerHisThrAlaAspIleSerSerHis	241
QY	667	TTGAGCAAAATGATGATTTTCAATGAGAAACAAAGCTTCAATGATGTGTTAGTG	726
Db	242	PheGlnGluAlaIleAspPheIleAspCysValaIrgLysGlyGlyLysValleuVal	261
QY	727	CAGCTTTTACCTGGGATCTCCCGCTCCGCAACCATGCTATGCGCTTACATGAAAGAG	786
Db	262	HisCysGluAlaGlyIleSerArgSerProThrIleCysMetAlaTyrlLeuMetLysThr	281
QY	787	ATGACATGCTCTTATAGATGAGACTTACAGATTGTGTGAATAAAGAAACCTTATATGT	846
Db	282	LysGlnPheArgLeuLysGluAlaPheAspTyrlIleLysGlnIArgIrgSerMetValSer	301
QY	847	CCAACTTCATTTTCTGCGGCCAACTCTGCACTATGAGAAAGATTGAAACAGACACT	906
Db	302	ProAsnPheGlyPheMetGlyGlnLeuLeuGlnIrgLysSerGlnIleLeuProSerThr	321
QY	907	-----GGAGCATCGGGCCCAAAAGCGAAATCTCAAGCTG	939
Db	322	ProAsnProGlnProProSerCysGlnGlyGluAlaIleGlySerSerLeuIleGlyHis	341
QY	940	CTGCACCTGGAGAAAGCCAAAT-----	960
Db	342	LeuGlnIrhLeuSerProAspMetGlnIleAlaTyrcysThrPheProAlaSerValleu	361
QY	961	GAACCTGTCTCTCTCTTCACAGGGGTGACAGAAACGAGACGCCCTCAGTCCACCC	1020
Db	362	AlaProValProThrHisSerThrValSerGluLeuSerAlaGserProValAla-----	379
QY	1021	TGTGCGCACTGCTGATCTCA 1041	
Db	380	-----ThrAlaThrSer 383	

RESULT 10
 T46405
 hypothetical protein DKFZp43401321.1 - human
 C/Species: Homo sapiens (man)
 C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 21-Jun-2002
 C/Accession: T46405
 R/Blum, H.; Baurerbach, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 Submitted to the Protein Sequence Database, January 2000
 A/Reference number: Z23034
 A/Accession: T46405
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-303 (AAA>
 A/Cross-references: EMBL:AL137704
 A/Experimental source: adult testis; clone DKFZp43401321
 A/Genetics:
 A/Note: DKFZp43401321.1

Alignment Scores:
 Pred. No.: 1,6e-13 Length: 303
 Score: 345.00 Matches: 101
 Percent Similarity: 48.54% Conservative: 49
 Best Local Similarity: 32.69% Mismatches: 109
 Query Match: 9.54% Indels: 50
 DB: 2 Gaps: 11

US-10-029-345A-108_COPY_538_2532 (1-1995) x T46405 (1-303)


```

Db      29 ValArgLAbProSerMet-AlaLeuGlyValSerGlnLeu-----41
Qy      326 GTAAACGTGGAAGAGAGCTTCACTCTGTTCACCTGTCGACGCTGGCTTTCGAGATTC 385
Db      42 ----AlaGlyArgSerArgCysLeuGlySerGlnSerGlnGlyGlyTyrGlnArgPhe 60
Qy      386 CTGCTGTTCCCTCGGCTCGTGAAGAAATCCACTCA-----G 427
Db      60 eTserGlnTyrProGlnPheCysSerGlyThrIlyAlaLeuAlaAlaIleProProPro 80
Qy      428 TCCCTACCTGCAATTTCTCAGCT-----TGC---TTCAGCTTGG 463
Db      80 alProProSerAlaThrGlnProLeuAspLeuGlyCysSerSerCysGlyThrProLeuH 100
Qy      464 CCAACATT---GGGCAACCCGGAATTTCTCCCAATCTTTATCTTGGCTGCACGAGATG 520
Db      100 IsArgGlnGlyGlyProValGlnIleLeuProPheLeuTyrLeuGlySerAlaTyrHis 120
Qy      521 TCCCTCAACAGAGCTGATACAGAGAAATGGATGGTATGCTTAAATCCAGCTATA 580
Db      120 laAlaArgArgAspMetLeuAspAlaLeuGlyIleThrAlaLeuLeuAsnValSerSer 140
Qy      581 CTTGTCCAAAGCCCTGATTTATCCCGAGTCTCATTTCTGCGTGCCTGTGAATGACA 640
Db      140 spCySerProAsn---HisPheGlnGlyHisIleTyrGlnTyrIleCysIleProValGlnAsp 159
Qy      641 GCTTTGTGAGAAATTTTCCGCTGTGTGACAAATCAGTAAATTTGATGAGAAAGCAA 700
Db      159 snHisIlyAlaAspIleSerSerThrPheMetGlnAlaIleGlnTyrIleAspAlaVal 179
Qy      701 AAGCTCCCAATGATGCTGTCTAGTACGCTTTAGCTGGATTCCTCCGCTCCGCCACA 760
Db      179 ysArgCysArgGlyArgValLeuValHisCysGlnAlaGlyIleSerArgSerAlaThr 199
Qy      761 TCCCTACCTGCTCATCATGATGAAGAGATGACATGCTTATGATGATGATGATGATTTG 820
Db      199 leCysLeuAlaIlyLeuMetMetIlyAspArgValaGlnLeuGlnGlnAlaPheGlnPhe 219
Qy      821 TGAAGAAAAGAGCCTATATATCTCCAAATTTCTGGGCAATCTCTGAGCT 880
Db      219 alIlyGlnArgArgSerIleIleSerProAsnPheSerPheMetGlyGlnLeuLeuGln 239
Qy      881 ATGAGAAAGATTAAGAACCAAGCTGAGCATGAGGCCAAGAGCAAACTCAAGCTGC 940
Db      239 heGlnSerGlnValLeuAlaThrSerCysAlaAlaGlnAlaIleSer-----254
Qy      941 TGCACCTGAGAGAGCCAAATGAACCTGCTCCCTGCTGTCTCAGAGGCTGACAGAAAGCG 1000
Db      255 -----ProSerGlyProLeu-----ArgGlnArgGlyL 264
Qy      1001 AGACGCCCTCAGTCCACCTGTGCCAGCTCTGTACTCTCAGAGCAGACAGCAAAAGC 1060
Db      264 yThrThrProAlaThrPro-----ThrSerGlnPheValPheSerPheP 278
Qy      1061 CCGTGCATCCCGCCAGGCTGCCAGCTGCCAGCTGCAGCCGCTGCTTGAAGACA 1120
Db      278 roVal---SerValGlyValHisSerAlaProSerSerLeuProTyrLeu-----His 295
Qy      1121 GCCCGCTGTAACAGGCGCTCAGT 1143
Db      295 eTProIleThrThrSerProSer 302

RESULT 11
T32494
hypothetical protein C05B10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32494
R:Geisels, C.; Wamley, P.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid C05B10.
A:Reference number: Z21178

```

```

A:Accession: T32494
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-365 <GB>
A:Cross-references: EMBL:AF036685; PIDN:AA88308.1; GSPDB:GN00022; CESP:C05B10.1
A:Experimental source: strain Bristol N2; clone C05B10
A:Genes: CESP:C05B10.1
A:Map position: 4
A:Insertions: 22/2, 46/2, 74/3, 163/2, 204/2, 248/1, 287/2

Alignment Scores:
Pred. No.: 1,93e-10 Length: 365
Score: 294.50 Matches: 83
Percent Similarity: 45.62% Conservative: 68
Best Local Similarity: 25.08% Mismatches: 137
Query Match: 8.15% Indels: 43
DB: 2 Gaps: 8

US-10-029-345A-108_COPY_538_2532 (1-1995) x T32494 (1-365)
Qy      19 GGAATCAAAATTTGTACTGAGAGCTTGTGCTCTCTGGAAGTGAACGGAAGATG 78
Db      12 GlyGlnGlnIleSerAlaGlnGlnPheAsnArgIlePheHisGln-----ArgAsnVal 29
Qy      79 CTGCTAATGATGATCCCGCCATTTGTGGAATACAAATACATCCACATTTTGAAGCATT 138
Db      30 IleValLeuAspCysArgSer-----AsnGlyAspSerValIlyArgAlaAsn 45
Qy      139 AATATCAACTGCTCCAGCTTATGAACGAGAGTGTCCACAGACAGCAAGTGTAAATTACA 198
Db      46 ArgLeuArgLeuProAlaLeuLeuGlnArgArgLeuMetGlyIleSerMetArgLeuSer 65
Qy      199 GAGCTCATCCAGCATTCACGGAACATTAAGCTTGACATTTGACATGACGAGAAAGTTGTA 258
Db      66 ThrValProAspLeuIlyAspLeuAsnAsnSerProAspGlnCysProGlnValLeuLeu 85
Qy      259 GTTTAGATCAAAAGCTCCCAAGT-----GTGCTCTCTCTCT 297
Db      86 IleProGlyAspSerIleGlnAspGlnGlnLeuSerAlaAlaLeuAlaArgAsnLeuLys 105
Qy      298 TCAGACTGTTTCTCACTGATCTTGTGGTAACTGAGAGAGCTTCACTCTGTTTAC 357
Db      106 SerAsnHisIlyArgHisIlePheValLeuGlyGlnProValIlyThrLeuLeuSerGlnPhe 125
Qy      358 CTGCTTGCAGAGTGGCTTGTCTGAGTTCT-----CGTTGTTTC 396
Db      126 ProThrLeuArgAspAlaAlaAspGlnAsnThrProAsnThrThrPheGlnMetAsnSerMet 145
Qy      397 CTTGGCTCTGTGAAGAAATTCACCTTAATGCTTCACTGATTTCTCAGCTTGTCTTA 456
Db      146 ProGlyGlnAlaSerIlyGlnGln-----AlaSerSerGlyProLeuLeu 160
Qy      457 CTGTGCGCAACATTTGGG-----474
Db      161 AsnLeuAsnGlnLeuArgLeuGlnGlyIlyAspGlnGlyIlySerGlnArgAlaGlnPhe 180
Qy      475 CCAACCGGAATTTCTCCAACTTTATCTTGGCTGCACGAGATGCTTCAACAGAG 534
Db      181 ProValIlyLeuThrAsnPheLeuTyrLeuGlyAsnAlaGlnThrAlaIlyAsnAsnArgAsp 200
Qy      535 CTGATACAGACAGAAATGGATTTGTATGTTTAAATCCAGCTATACCTCTCCAAAGCCT 594
Db      201 ValLeuIlyAspIlySerIleSerHisValIleAsnValThrSerAsnLeuProAsnThr 220
Qy      595 ---GACTTATCCCGAGTCTCATTTCTGCGTGCCTGTGATGATGACAGCTTTGTGAG 651
Db      221 PheGlnGlnAspProAsnMetArgTyrLeuAlaGlyIleSerAlaAspAsnAsnIleSerHis 240
Qy      652 AAAATTTCCGCTGTGACAAATCAGTAAATTTGATGAGAAAGCAAAAGCTTCAAT 711
Db      241 AsnLeuThrIlyPhePheProGlnAlaIleSerPheIleAspAspAlaArgAsnAsp 260

```



```

QY      712  GATGATGTTCTAGTGCAGCTGTTTAACTGGAGATCTCCGTCACCAATGCTATGCGC 771
      |||
Db      261  SerAlaCysLeuValHisCysLeuAlaGlyIleSerArgSerValThrIleCysLeuAla 280
QY      772  TACATCATGAAAGAGATGAGACATGCTTTAGATGAGACTTACAGATTTGTGAAGAAAA 831
      |||
Db      281  TTTLeuMetLysThrGluMetCysThrLeuAspSerAlaTyrGluTrpValGlnLysArg 300
QY      832  AGACCTATATATCTCCAAACTTATTTCTGGGCGCAATCTCGGACTATGAGAAAG 891
      |||
Db      301  AsnAlaSerIleAlaProAsnIleHisPheMetGlyGlnLeuThrAspTyrGluLysMet 320
QY      892  ATT-----AAGAACCAAGCTGAGACATCA 915
      |||
Db      321  LeuGlyLeuAsnSerAsnArgValGlyValSer 331

RESULT 12
T18915
hypothetical protein C04F12.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18915
R: Lloyd, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19044
A:Accession: T18915
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-272 <WIL>
A:Cross-references: EMBL:Z81461; PIDN: CAB03837.1; GSPDB: GN00019; CESP: C04F12.8
C:Genetics:
A:Gene: CESP: C04F12.8
A:Map position: 1
A:Insertions: 53/1; 84/3; 204/2

Alignment Scores:
Pred. No.: 4,256-07 Length: 272
Score: 239.50 Matches: 77
Percent Similarity: 47.84% Conservative: 45
Best Local Similarity: 30.20% Mismatches: 110
Query Match: 6.63% Indels: 23
DB: Gaps: 8

US-10-029-345A-108_COPY_538_2532 (1-1995) x T18915 (1-272)
QY      466  AACATGAGG-----CCAAACCAATTTCTCCCAATCTTATCTTGCTGCCAG 513
      |||
Db      4  AsnValGlyValLeuGlyGlnMetSerGlnIleAsnAspHisLeuTyrLeuSer---Gly 22
QY      514  CGAGATGCTCTTAACAAAGAGCTGATAGACAGAAATGGATTTGGTTATGTTAAATGCC 573
      |||
Db      23  AlagIValLeuLysProAspLysIleLysGlnArgLysIleAsnMetIleValAsnAla 42
QY      574  AGCTTAACCTGTCGCAAGCTGATTTATCCCGAGTCTATTTCTGCGTGCCTGTG 633
      |||
Db      43  ThrThrGluGluProSer---ThrTyrMetGlnGlyValAspThrMetLysIleArgIle 61
QY      634  AATGACAGCTTTTGTGAGAAATTTTGGCTGTGGCAAAATGAGATTTTCAATTTGATTTAG 693
      |||
Db      62  GluAspHisProTyrAlaArgLeuAsnGlnHisPheAspValAlaAlaAspLysIleArg 81
QY      694  AAGAGAAAGCTCCAAATGATGATGTTCTTATGCTGATGCTGTTAGCTGGATCTCCGCTCC 753
      |||
Db      82  AsnValLysGluArgGlyGlyLysThrLeuValHisCysMetAlaGlyValSerArgSer 101
QY      754  GCCAACCATGCTATGCTGCTCATCATGATGAGAGATGAGATGCTGCTTATGATGAGATTTCC 813
      |||
Db      102  AlaSerLeuValMetIleTyrLeuValLysGlnHisMetThrLeuLysArgGlnAlaTyr 121
QY      814  AGATTTGGAAGAAAAAGACCTACTATATCTCCAAATTTCAATTTTCTGGGCGCAATCTC 873
      |||
Db      122  HisTyrValLysAlaAlaArgProIleIleArgProAsnValGlyPheThrLysGlnMet 141

```

```

QY      874  CTGACCTATGAGAAAGAAATTAAAGAACAGATGAGATCAGGCGCAAGCAAACTC 933
      |||
Db      142  ValAspTyrGluLysArgLeuArgGly-----ThrAlaSerVal 154
QY      934  AAGCTGTCACCTTGAGAACCCAAATGAACTTGTCCCTGCTGCTCAGAGGGTGA--- 990
      |||
Db      155  LysMetValGlnThrProGluCysAspMetProIleProAspValTyrAlaAspAla 174
QY      991  -----CAGAAAAGCAGACGCCCTTCACTCACCCTGTCGCACTCTGTACTCAGAG 1044
      |||
Db      175  ArgArgMetGlnIleAsnArgGluIleSerArgHisSerLeuProSerAlaGlnSerSer 194
QY      1045  GCAGCAGGACCAAGGCCCGTGCATCCCGCAGGGT-----CCGAGCTGCGCCAGCTGCAG 1101
      |||
Db      195  AlaSerLysGlnArgAlaPheSerAlaSerThrTyrArgProSerValSerAla--- 213
QY      1102  CCGTGCCTTTAGAGACAGCCCGCTGTGATGAGCGCTCAGTGGCTGACCTGTCCGA 1161
      |||
Db      214  ---SerSerLeuAlaThrSerSerLeuArgArgAlaTyrSerProSerSerLeuProAla 232
QY      1162  GACAGCTGGAAGACAGCAATTAAGCTCAAGCGTCTCTCTCTG 1206
      |||
Db      233  SerSerLeu-----AlaLeuThrThrSerTyrSerLeu 243

RESULT 13
T21380
hypothetical protein F26A3.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21380
R:McMurray, A.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19415
A:Accession: T21380
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-226 <WIL>
A:Cross-references: EMBL:Z78419; PIDN: CAB01700.1; GSPDB: GN00019; CESP: F26A3.4
C:Genetics:
A:Gene: CESP: F26A3.4
A:Map position: 1
A:Insertions: 117/2; 150/3; 186/3

Alignment Scores:
Pred. No.: 9,776-07 Length: 226
Score: 233.50 Matches: 64
Percent Similarity: 49.55% Conservative: 47
Best Local Similarity: 28.57% Mismatches: 80
Query Match: 6.46% Indels: 33
DB: Gaps: 6

US-10-029-345A-108_COPY_538_2532 (1-1995) x T21380 (1-226)
QY      478  ACCCGAATTTCTCCCAATTTTATCTTGCTGCGCAGAGATGCTCTCAACAGAGCTG 537
      |||
Db      15  SerGlnLeuValProGluLeuPheIle---CysGlyValSerAlaLeuSerLysAspGlu 33
QY      538  ATACAGCAGAAATGAGATGTTATGTTTAAATGCGCACTATTCCTGTCCAAAGCTGAC 597
      |||
Db      34  MetLysLysHisLysLysIleThrHisIleAsnAlaThrThrGluValProAsnLeuArg 53
QY      598  TTTATCCCGAGTCTCATTTCTGCTGCTGCTGCTGATGATGACAGCTTTGTGAGAAATT 657
      |||
Db      54  SerLeuLysLysPheGlnArgThrLysLeuTyrPheGlnAspThrProGlnThrTyrIle 73
QY      658  TTGCGCGGTGAGCAAAATCAGATTTTCAATTTGAGAAAGAAAGCTTCAATGAGATGT 717
      |||
Db      74  TyrProHisLeuGluLeuGlnSerAspGlnIleGlnAlaLeuIleAlaAspGlyGlyLys 93
QY      718  GTTCTAGTGCAGTTTTATGCTGAGATCTCCGCTCCGCAACATGCTATGCGCTACATC 777
      |||

```

Db 94 ValIeuValHicCyValaIaglyValIserArgSerAlaSerIleCysLeuAlaPheLeu 113
QY 778 ATGAAGAGATGACATGCTTTAGATGAAGCTTACGATTTTGAAAGAAAAGACCT 837
Db 114 LeuIySerYrArgCysArgAenLeuArgIuAlaIyRhiIleuMetIySerIySer 133
QY 838 ACATATCTCCAACTTCAATTTCTGGCCAACTCTGTGACATATGAGAAAGATTAAG 897
Db 134 MeValArgProAenLeuGlyPheTrpArgIleuIleAlaIyRgiIuAenValIyR 153
QY 898 AACCAAGCTGAGACATCAGGAGCCAAAGACAACTCAAGCTGTGACCTGAG--AG 954
Db 154 GluAenAlaGlySer-----ValArgIeuValArgAenGluAlaIagIn 167
QY 955 CCAATGACCTGTCTCTGTCTCAGAGGGTGCAGACAAAGAGAGAGCCCTCAGT 1014
Db 168 ProGluInIleuLeuPro----- 173
QY 1015 CCAACCTGTGCGCTGTACCTCAGAGCAGACAGACAAAGAGCCGTGATCCGCC 1074
Db 174 -----AspValIyRLeuAenIleAlaIle-----ProAlaArgProAla 186
QY 1075 AGCGTGCCAGCGTGCACGCTGCAGCGTGCCTGTAGAGACAGCCGCTGTACAG 1134
Db 187 Ser-----ProGluInAenProAenMetIleProAenGluProArgGluArg 202
QY 1135 GCGCTCAGTGGG 1146
Db 203 ArgAenSerGly 206
RESULT 14
T16056
hypothetical protein F13D11.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001
C/Accession: T16056
R/Fulton, L.
submitted to the EMBL Data Library, November 1995
A/Description: The sequence of C. elegans cosmid F13D11.
A/Reference number: S69020
A/Accession: T16056
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-186 <PUL>
A/Cross-references: EMBL:U40939; NID:G1072175; PID:G1072176; PIDN:AAA81700.1; CESP:F13D11.3
A/Gene: CESP:F13D11.3
A/Introns: 30/3; 57/3; 85/3; 125/3; 172/3
Alignment Scores:
Pred. No.: 2,97e-06 Length: 186
Score: 225.50 Matches: 57
Percent Similarity: 52.60% Conservative: 34
Best Local Similarity: 32.95% Mismatches: 59
Query Match: 6.24% Indels: 23
DB: 2 Gaps: 5
US-10-029-345A-108_COPY_538_2532 (1-1995) x T16056 (1-186)
QY 478 ACCGGAATCTTCGCAATCTTTATCTT-----GGTCCAGCAGAGATGCTCAAC 528
Db 12 ThGInValArgProAenLeuArgIuAlaIyRgiIyCys-----IleThr 27
QY 529 AAGCACTGATACAGCAAGATGGATTGTATGTGTTAAATGCCACTATACCTGTCCA 588
Db 28 ProSerLeuLeuYsgInIyRhiIleThRhiGlyValAenPheCysThrAenLeuYthr 47
QY 589 AAGCGTACTTTATCCCGAGTGCATTTCTGCGTGTGCTGTGATGACAGCTTTGT 648
Db 48 LysPro-----IleIySgIyLeuAenPArgIleIuValProValAenAenThrLeu 65
QY 649 GAGAAATTTTGGCGTGTGACAAATCATGATGATTTCAATGAGAAGCAAGAGCTCC 708
||||| : : : : : ||| : : : : : |||

Db 66 AlaIyIleThrGInIyRPhGluProValIalYyRiIleGluAenAlaIyGIn 85
QY 709 AATGATGTCTTACGTGACCTGTATGCTGGATCTCCCGCTCCGACCATCGTATC 768
Db 86 GlyHiaAenThrValIyRyCysAlaIaglyValIserArgSerAlaThrIleThr 105
QY 769 GCGTACATCAGAGAGATGACATGCTTTAGATGAAGCTTACAGATTGTGAAGAA 828
Db 106 ValIyRLeuMetValIyRgIuAenLeuSerIleGluAlaIyRLeuGInValAenGIn 125
QY 829 AAAAGACTATATCTCCAACTTCAATTTCTGGCCAACTCTGACCTATGAGAG 888
Db 126 ValArgProIleIleSerProAenIleGlyPheTrpArgIleuMetIleAenPheGluIyR 145
QY 889 AAGATTAGAACCAAGACTGAGCA-----TCAGGCGCAAGCAAACTCAACTG 939
Db 146 GluArgAenGlyAenAlaIserValGluLeuIleSerIyArgMetAlaArg----- 162
QY 940 CTGACCTGAGAGAGCAATGAACTGCTGCTGTGTC 978
Db 163 -----ProValProSerVal 167
RESULT 15
T39698
protein tyrosine phosphatase - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C/Accession: T39698
R/Wood, V.; Skelton, J.; Churcher, C.M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, July 1999
A/Reference number: Z21870
A/Accession: T39698
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-330 <MO>
A/Cross-references: EMBL:AL109652; PIDN:CA851765.1; GSPDB:GN00067
A/Experimental source: strain 972h-; cosmid c17A3
C/Genetic:
A/Gene: p1041
A/Map position: 2
Alignment Scores:
Pred. No.: 4,05e-06 Length: 330
Score: 223.50 Matches: 53
Percent Similarity: 52.59% Conservative: 18
Best Local Similarity: 39.26% Mismatches: 63
Query Match: 6.18% Indels: 1
DB: 2 Gaps: 1
US-10-029-345A-108_COPY_538_2532 (1-1995) x T39698 (1-330)
QY 478 ACCGGAATCTTCGCAATCTTTATCTTGGCTGCAGCAGAGATGCTCAAGAGACTG 537
Db 48 SerGluIleSerIyAenLeuYrIleSerSerTrpArgThrAlaSerGluLeuValSer 67
QY 538 ATACAGCAAGATGGATTGTATGTGTTAAATGCCAGCTATACCTGCCAAGCCCTGAC 597
Db 68 ThrSerAspIySgIyIleAepIyRhiLeuSerAlaMetSerIleAenProAenLeuSer 87
QY 598 TTTATCCCGAGCTCATTTCTGCGTGTGCTGTGATGACAGCTTTGTGAGAAAT 657
Db 88 ---ValProGluInIleGInIleSerIyRLeuTrpLeuGInIleGluAenPheSerSerGInAenIle 106
QY 658 TTGCGGTGTGAGCAATATGATGATTTGATGAGAAAGCAAGCTCCATGATGT 717
Db 107 LeuGInIyRPhGluIySerAenIySgIyRhiAlaPheAlaIleuSerIyAenAlaIyS 126
QY 718 GTTCTAGTGCATGTTTACGTGGATCTCCGCTCCGACCATCGTATCGCTCATC 777
Db 127 ValIeuValHicCySgIyAenAlaIyIleSerArgSerValIyRLeuValAlaIyRLeu 146
QY 778 ATGAAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAGCT 837
||||| : : : : : ||| : : : : : |||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:21 / Search time 20.9747 Seconds

(Without alignments)
9905.280 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 3615
Sequence: 1 atggcccatgagatgctg.....tggaatcatgagctcc 1995

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlp
-O=/cgr21/USPTO.spool.p/US10029345/runat_21062004_122815_4177/arp_query.fasta_1.3278
-DB=SwissProt 42 -OFMT=fasta -SUFFIX=esp -MINMATCH=0.1 -IOCPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCNALGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=exc -HEAPSIZE=500 -MINLEN=50 -MAXLEN=200000000
-USBR=US10029345_@CGC_1.1_33_@runat_21062004_122815_4177 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt 42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3406	94.2	665	DUSG_HUMAN	Q9by84 homo sapien
2	1326	36.7	625	DUS8_HUMAN	Q13202 homo sapien
3	1302	36.0	663	DUS8_MOUSE	O09112 mus musculu
4	515	14.2	619	VHPI_CAEEL	Q10038 caenorhabdi
5	471.5	13.0	381	DUS6_HUMAN	Q16828 homo sapien
6	469	13.0	482	DUS6_HUMAN	Q95666 homo sapien
7	468	12.9	483	DUS6_MOUSE	Q96880 mus musculu
8	466.5	12.9	381	DUS6_MOUSE	Q96880 mus musculu
9	465.5	12.9	381	DUS6_MOUSE	Q96880 mus musculu
10	447	12.4	394	DUS4_HUMAN	Q13115 homo sapien
11	444	12.3	375	DUS4_CHICK	Q97471 gallus galli
12	439.5	12.2	395	DUS4_RAT	Q62767 rattus norv
13	433	12.0	367	DUS1_MOUSE	P28563 mus musculu
14	423	11.7	367	DUS1_RAT	Q64623 rattus norv
15	420	11.6	367	DUS1_HUMAN	P28562 homo sapien
16	418	11.6	314	DUS2_HUMAN	O05923 homo sapien
17	417	11.5	384	DUS9_HUMAN	Q99956 homo sapien
18	409	11.3	318	DUS2_MOUSE	Q05922 mus musculu

19	390	10.8	320	DUS7_HUMAN	Q16829 homo sapien
20	386.5	10.7	320	DUS7_MOUSE	Q91246 mus musculu
21	367	10.2	384	DUS7_RAT	O54838 rattus norv
22	361.5	10.0	280	DUS7_RAT	O63340 rattus norv
23	360.5	10.0	384	DUS7_HUMAN	O16650 homo sapien
24	269	7.4	198	DUSE_HUMAN	O95147 mus sapien
25	268	7.4	198	DUSE_MOUSE	O95147 mus musculu
26	205.5	5.7	188	DUS1_HUMAN	O8ne30 homo sapien
27	205.5	5.7	217	DUS1_HUMAN	O8ne30 homo sapien
28	205.5	5.7	223	DUS1_HUMAN	O8ne30 homo sapien
29	205	5.7	223	STYX_MOUSE	O60959 mus musculu
30	204	5.6	278	PWPI_SCHPO	O13453 schizosacch
31	202	5.6	364	FPM1_YEAST	Q02256 saccharomyc
32	201.5	5.6	313	MSTY_HUMAN	O95618 homo sapien
33	195.5	5.4	185	DUS3_HUMAN	P51452 homo sapien
34	195.5	5.4	225	DUSE_HUMAN	O9112 homo sapien
35	194.5	5.4	185	DUS3_MOUSE	Q9d733 mus musculu
36	189	5.2	489	MSG5_YEAST	P38530 saccharomyc
37	185	5.1	807	YBPT_YEAST	P38148 saccharomyc
38	183.5	5.1	276	PTP3_CHLEU	O39491 chlamydomon
39	177	4.9	125	DUSE_MOUSE	O8r42 mus musculu
40	173.5	4.8	339	DUSC_MOUSE	Q9d022 mus musculu
41	170.5	4.7	292	YO42_CAEEL	P34650 caenorhabdi
42	163.5	4.5	340	DUSC_HUMAN	O9un16 homo sapien
43	161	4.5	198	DUSD_HUMAN	O9un16 homo sapien
44	158.5	4.4	1229	P121_HUMAN	O9y2n3 homo sapien
45	157	4.3	209	YIL3_YEAST	P40479 saccharomyc

ALIGNMENTS

RESULT 1
DUSG_HUMAN
ID DUSG_HUMAN STANDARD; PRT; 665 AA.

AC Q9by84; Q9CG03;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 16 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 7) (MAP kinase
DE phosphatase 7) (MKP-7).
GN DUSP16 OR MKP7 OR KIAA1700.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21486429; Pubmed=11489891;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "MKP-7, a novel mitogen-activated protein kinase phosphatase,
RT functions as a shuttle protein.";
RT J. Biol. Chem. 276:39002-39011(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21082932; Pubmed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:347-355(2000).
CC - FUNCTION: Involved in the inactivation of MAP kinases.
CC - CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC - CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC - CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC - SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC - SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC - SIMILARITY: Non-receptor class dual specificity subfamily.
CC - SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB052156; BAB40814.1; -
 DR EMBL; AB051487; BAB21791.1; ALT_INIT.
 DR HSSP; O16828; IMKP.
 DR Genew; HGNC:17909; DUSP16.
 DR MIM; 607175; -
 DR GO; GO:0005737; Cytoplasm; TAS.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0004721; F:protein phosphatase activity; TAS.
 DR GO; GO:0016311; P:dephosphorylation; TAS.
 DR GO; GO:000188; P:inactivation of MAPK; TAS.
 DR GO; GO:0045209; P:leptomycin B-sensitive MAPK phosphatase nuc. .; TAS.
 DR GO; GO:0045204; P:MAPK nucleus export; TAS.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; MAPKPHNTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANESE_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR HydroLase; Nucleic protein.
 FT DOMAIN 22 137 RHODANESE.
 FT 227 289 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 244 244 PHOSPHOCYSTEINE INTERMEDIATE (BY
 FT SEQUENCE 665 AA; 73101 MW; 1BD853FE08460DF CRC64;
 SIMILARITY).

Alignment Scores:
 Pred. No.: 3 66-206
 Score: 3406.00
 Percent Similarity: 99.85%
 Best Local Similarity: 99.70%
 Query Match: 94.22%
 DB: 1
 Gaps: 0

US-10-029-345a-108_COPY_538_2532 (1-1995) x DUSG_HUMAN (1-665)

QY 1 ATGGCCCATGAGATGTTGAACTCAATTTGTTACTGAGAGTTGGTGGCTTGCTGGAA 60
 DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
 QY 61 AGTGAACGGAAGAGTGTCTATTTAGTACCGGCGCATTTGGGAATCAATACATCC 120
 DB 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 QY 121 CACATTTTGAAGCCATTATATCACTCTCAAGCTTATGAAGCGAAGTTGCAACG 180
 DB 41 HisIleuGluAlaIleAsnIleAsnGlySerIleuMetIleAsnGluGln 60
 QY 181 GACAAAGTGAATTAATTAAGAGCTCATCCAGATTACCGAAACATTAAGTTGAT 240
 DB 61 AspIleValLeuIleThrGluLeuIleGlnHisSerIleAlaIleValAspIleAsp 80
 QY 241 TGCAGTCGAAGGTTGATGTTTACGATCAAGCTCCCAAGATTTGCTCTCTTCA 300
 DB 81 CysSerGlnIleValIleValIleValIleValIleValIleValIleValIleVal 100
 QY 301 GACTGTTTCTCACTGACTTCTGGGTAACCTGGAAGAAGCTTCACTGTTGACCTG 360
 DB 101 AspCysIleuThrValLeuLeuGlyIleValIleValIleValIleValIleValIle 120

QY 361 CTTCGAGTGGGTTTCTGAGTTCTCTGTTGTTTCCCTGGGCTCTGGAAGAAATCC 420
 DB 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGluGlySer 140
 QY 421 ACTCTAGTCCCTACCGGATTTCTGAGCTTGCTTACTGTTGCCAATTTGGCCAAAC 480
 DB 141 ThrIleValProThrCysIleSerGlnProCysIleuProValAlaAsnIleGlyProThr 160
 QY 481 CGAATTTCCCAATCTTTATCTTGAGCTCCAGGAGATGCTCCCAAGAGGCTGATA 540
 DB 161 ArgIleuProAsnIleuTyrLeuGlyCysGlnArgAspIleuAsnIleGluMet 180
 QY 541 CAGCAGAAATGGAATGTTAGTTATGTTAAATCCAGCTTAACTGTTCCAAAGCTGACTT 600
 DB 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProIleProAspHe 200
 QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGTAATGACAGCTTTTGTGAGAAATTTTG 660
 DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIleLeu 220
 QY 661 CCGTGTGGAGCAAAATGATGATTTTATTTGAGAAAGCAAAAGCTTCAATGATGTT 720
 DB 221 ProThrLeuAspIleSerValAspPheIleGluValAlaIleValAsnGlyCysVal 240
 QY 721 CTAGTCACTGTTTAACTGAGATCTCCGCTCCGACACATGCTATGCTTACATG 780
 DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
 QY 781 AAGAGATGAGACATGCTTAAATGAGCTTCAATTTGTAAGAAAGAAAGAAAGCTTACT 840
 DB 261 LysArgMetCysPheSerLeuAspGluAlaTyrArgPheValGlyGluValArgProThr 280
 QY 841 AATCTTCAAACTTCAATTTCTGGGCAACTCTGTAATGAGAAAGAAATTAAGAAC 900
 DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuAspIleTyrGluValValIleVal 920
 QY 901 CAGATGAGACATCAAGGCGCAAGCAAAAGCTCAAGCTGCTGCACTTGAGAAAGCAAT 960
 DB 301 GlnThrGlyAlaSerGlyProIleSerIleuIleuIleuIleuIleuIleuIleuIleu 320
 QY 961 GAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 DB 321 GluProValProAlaValSerGluGlyGlyGlnIleuSerGlnIleuIleuIleuIleu 340
 QY 1021 TGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
 DB 341 CysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSerVal 360
 QY 1081 CCGAGCTGCGCCAGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 DB 361 ProSerValProSerValGlnProSerLeuLeuGluIleuSerProLeuValGlnAlaLeu 380
 QY 1141 AGTGGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 DB 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAlaIleuValIleuValSerPhe 400
 QY 1201 TCTGTGATATCAATGATTTATATTAATGAGAGATGAGAGATGCTTAAATGCTTC 1260
 DB 401 SerLeuAspIleuSerValSerIleuSerAlaSerMetAlaSerIleuHisGlyPhe 420
 QY 1261 TCCCATCAAGAAATGCTTGGAAATCAAAACCTTCCACTACTGATGGAAGCAAC 1320
 DB 421 SerSerSerIleuAspAlaLeuGluTyrTyrIleuProSerThrThrLeuAspGlyThrAsn 440
 QY 1321 AAGCTATGCAAGTTTCCCTGTTGAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 DB 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
 QY 1381 GATTAAGAGAGAGAGAGAGAGATCCCAAGAGCTGAGACCGGCAAGGCTTGAAGAGCAG 1440
 DB 461 AspIleGluGluAlaSerIleProIleValIleuGlnThrAlaArgProSerAspSerGln 480
 QY 1441 AGCAAGCATTCATTCGCTGAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500

Db	Accession	Protein Name	Length
Db	481	SerlysbAglyLeuHhSerValArgThrserserSerGlyThAlaGlnArgserLeuLeu	500
Qy	1501	TCCTCACTGATCGAAGTGGAGCGCTGGAGACAATTACACACAGCTTCCTTTGGC	1560
Db	501	SerProleuHhArgserGlySerValGlnAparantYHhIthSerPheLeuPheGly	520
Qy	1561	CTTTCACACGACGACGACACTTCACGAAAGTGTGCTGGCTTGAGGGCTGGCAC	1620
Db	521	LeuSerThrserGlnGlnHhIbLeuThrlybSerAlaGlyLeuGlyLeuylbGlyTrpHhIs	540
Qy	1621	TCGGATATCTGGAGCCCGACAGCTCTACCCCTCCCTGACACAGAGCTGATATTGGC	1680
Db	541	SerapHhIleLeuAlaProGlnThrserrThrProserLeuThrsSerTrpYrPheAla	560
Qy	1681	ACAGAGCTCTGACACTTCTACTCTGCTCCCTCAGCCATCTACGAGGACAGTGCAGTTACTCT	1740
Db	561	ThrGlnSerSerHhIbPheTybSerAlaSerAlaIleTybGlyGlySerAlaSerTybSer	580
Qy	1741	GCTTACAGCTGACGACGCTGCTCCACTTGGCGGACCAAGTCTATTCTTGCGCAGGCGG	1800
Db	581	AlaTybSerTybSerGlnLeuProThrTybGlyAlaPbGlnValTybSerValAlaArgArg	600
Qy	1801	CAGAAAGCAAGTACAGAGCTGACTCGCGCGGAGCTGGSCATGAGAAGAGCCCTTTGAA	1860
Db	601	GlnTybProserAspArgAlaAspSerAlaArgSerTrpHhIbGlnIbSerProPheGln	620
Qy	1861	AAGCAGTTTAAACGACGAAGTGCACAAATTGAAATTGGAGAGCATCATGTTCAGAGAAC	1920
Db	621	LybGlnPheTybArgArgTybSerGlybGlnMetGlnPheGlyGlybSerIleMetserGlnuAn	640
Qy	1921	AGGTCACGGGAGAGACCTGGGGGAAAGTGGGAGCTGACTGACTTTTGGGAGCATGAA	1980
Db	641	ArgSerArgGlnGlnLeuGlnGlybValGlybSerGlnSerSerPheSerGlybMetGln	660
Qy	1981	ATCATTTGAGCTCTCC 1995	
Db	661	IleIleGlnValSer 665	

RESULT 2

DUS8 HUMAN

ID DUS8 HUMAN STANDARD; PRT; 625 AA.

AC Q13202;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)

DE (Dual specificity protein phosphatase hnh-5).

GN DUSP8 OR VHS.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_Taxid:9606;

XX [1]

RN SEQUENCE FROM N.A.

RP TISSUE=Fetal brain;

RC MEDLINE=96009533; PubMed=7561881;

RX Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.,

RT "hnh-5: a protein tyrosine phosphatase abundant in brain that

RT inactivates mitogen-activated protein kinase.";

RL J. Neurochem. 65:1823-1833(1995).

CC -1 FUNCTION: This protein shows both activity toward tyrosine-protein

CC phosphatase as well as with serine/threonine-protein phosphatase (By

CC similarity).

CC -1 CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein

CC -1 tyrosine + phosphate.

CC -1 CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +

CC phosphate.

CC -1 TISSUE SPECIFICITY: Abundant in brain, heart and skeletal muscle.

CC -1 SIMILARITY: Belongs to the protein-tyrosine phosphatase family.

CC Non-receptor class dual specificity subfamily.

CC -1 SIMILARITY: Contains 1 rhodanese domain.

CC -----

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC      EMBL; U27193; AAA83151.1; -.
CC      DR HSSP; Q16828; IMKP.
CC      DR Genew; HGNC:3074; DUSP8.
CC      DR MIM; 602038; -.
DR   GO: 00005737; C:cytoslaem; TAS.
DR   GO: 00004725; F:protein tyrosine phosphatase activity; TAS.
DR   GO: 0000188; P:inactivation of MAPK; TAS.
DR   GO: 0006470; P:protein amino acid dephosphorylation; TAS.
DR   InterPro: IPR000340; PS phosphatase.
DR   InterPro: IPR008343; MAPK phosph.
DR   InterPro: IPR001763; Rhodanese-like.
DR   InterPro: IPR000387; TYR_phosphataae.
DR   Pfam; PF00782; DSPc; 1.
DR   PRINTS; PR01764; MAEPKPHPTASE.
DR   SMART; SM00195; DSPc; 1.
DR   SMART; SMO0450; RHOD; 1.
DR   PROSITE; PSS0206; RHODANSE_3; 1.
DR   PROSITE; PSS0383; TYR_PHOSPHATASE_1; 1..
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN 23    138        RHODANSE.
FT   DOMAIN 162   430        PROTEIN-TYROSINE PHOSPHATASE.
FT   DOMAIN 310   550        PRO-RICH.
FT   ACT_SITE 246   246        PHOSPHOCYCSTEINE INTERMEDIATE (BY SIMILARITY).
SQ     SEQUENCE 625 AA; 65840 MW; DCBEA14487219666 CRC64; .

Alignment Scores:
Pred. No.:          8.94e-76           Length:         625
Score:             1326.00            Matches:         314
Percent Similarity: 58.54%            Conservative:    94
Best Local Similarity: 45.05%          Mistmatches:    185
Query Match:       36.68%              Indels:         104
DB:                1                  Gaps:           21

US-10-029-345A-108_COPY_538_2532 (1-1995) x DUSP_HUMAN (1-625)
QY      1 ATGGCCCATGATGATGATTGAACAACCAATTGT--ACTGAGAGTTGGTGCTCTGCTG 57
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      1 MetAlaIgllybpaIrglnePrOalrGlValMetncaphylalysIySleuAlasertleueu 20
QY      58 GAAAGTGAACGCgAAAAAGTCGCTCCTAATTGATnAGCGGCCCATTTGTtGAATAACA 117
      ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      21 ArgelIyprroGIlyglYproLeuvalIllelpserlnxgrSerPrevalIGlutYAmser 40
QY      118 TCCCAcATTTTGGAAccCAATTAAATAcATCAcTgcTTCcAAcGTTTAAGAacCGAAgTTGCA 177
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      41 TrpHlsVallleuserserValasnIllCyseSyserIysleuValIsyrGrIleugIn 60
QY      178 CAGAACAAAgCTTAATTATTAcGACCtcATCCAGAtTCAGGCAAcCTTAAGtgTAGAcATT 237
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      61 GinelyIybaIVmTrllealagIndeuIIegInProAlaaIaaySeGINvaIgIuaLa 80
QY      238 GATTCAGAcTCAGAAgGTtTGAATTAGcATAAAGctCCCAAGATgtTGcCTCTCTCT 297
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      81 ThrGIuPrOGInapervalValvalIalyAsplnserrThrarGaPaIalaservalleuaLa 100
QY      298 TCAGAcGTttTYTTCACtGNActTCCTTGcgTRAAcTGAGAAgAGcTTCAACTcGTTCAC 357
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      101 AlaApSerPreleuSerIlleuleuSerIysleuaspGLyCySPheASPserValala 120
QY      358 CTGCTTGAGAGTGGGgTTTGCTGcAGtTTCTCTGgTTTTCCTGcGCTCTGcGAAGANA 417
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      121 lleuThrngIyglYpHeaIathPheserSerCySphePrOglyneCYseluGIylys 140
```

```

QY 418 ---TCCAGTCTAGTCCCTTACTGCAATTTCTCAGGCTTGTCTTACTGCTGTGGCAATTTGGG 474
Db 141 ProAlaAlaLeuLeuProMetSerLeuSerClnProCysLeuProValProSerValGly 160
QY 475 CCAACCCGAATTTCTTCCCAATCTTATCTTGGCTGGCAGAGATGCTCTCAACAAGAG 534
Db 161 LeuThrArgIleLeuProHisIleuArgIleuGlySerClnIlyAspValIleuAsnIlyAsp 180
QY 535 CTGATACAGCAAGATGGGATTTGGTATGTGTAAATGCGAGTATCTGTCCAAAGCT 594
Db 181 LeuThrArgIleuGlnGlyIleSerIlyValIleuAsnAlaSerAsnSerCysProIlyPro 200
QY 595 GACTTATATCCCGAGTCTCATTTCTGCGNGGCTGTGAATGACAGCTTTTGGAGAA 654
Db 201 AspPheIleCysGlnSerArgPheMetArgValProIleAsnAspAsnIlyCysGlnIly 220
QY 655 ATTTGGCCGNGTGGCAAAATCAGATGATTTTCAATGAGAAAGCAAAAGCTCCAGATGGA 714
Db 221 LeuLeuProThrLeuAspIlySerIleGlnPheIleAspIlyAlaIlyLeuSerSerCys 240
QY 715 TGTGTTCTAGTCACTGTTAGTGGGATCTCCGCTCCGCACTGCTATGCTTAC 774
Db 241 GlnValIleValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIlyr 260
QY 775 ATCATGAAGAGATGAGCATGTTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAA 834
Db 261 IlemelIyStrhMetGlyMetSerSerAspAlaIlyrArgPheValIlysaPaIyArg 280
QY 835 CCTATATATCTCAAACTTCAATTTTCTGGCCCACTCTGAGCTATGGAAGAAAGATT 894
Db 281 ProSerIleSerProAsnPheAsnIleuGlnIlyGlnIleuGlnIlyrGlnIlyrGlnIlyr 300
QY 895 AAGAACGAGCATGAGCATCAGGCCCAAGAGCAAACTCAAGCTGCTGACCTGGAGAG 954
Db 301 LysLeuLeuAlaAlaLeuGlnIlyAspProGly-----Thr 312
QY 955 CCAAT-----GAACCTGTCCCTGCTGTCTGACAGAGCTGAGACAGAAAGCAGAGC 1005
Db 313 ProSerGlyThrProGlnIlyProProProSerProAlaAlaGlyAlaProLeuProArgLeu 332
QY 1006 CCCCTGAGTCCACCTGTGCGCACTGCTGAC-----TCAGAGCCAGAGCAACAAG 1059
Db 333 Pro-----ProProThrSerGlnSerIleAlaIlyrThrGlyAlaAlaIlyrGlnIly 350
QY 1060 CCCGTGATCCCGCAGCGTCCCGCAGCGTCCCGCAGCGTCCCGCTGTAGAGAGC 1119
Db 351 GlyLeuSerAlaGlyIlyGlnIlyProProAlaProProThrProProAla-----Thr 367
QY 1120 AGCCGCTGTACAGCGCTCAGTGGGCTGACCTGTCCGAGACAGAGCTGGAAGAGC 1179
Db 368 SerAlaLeuGlnIleGlnIlyLeuArgIlyLeuHisLeuSerSerAlaPheGlnIlyAspThr 387
QY 1180 AATAAGCTCAAGCGTCTCTCTCGATGATCAATCAATGTTTCAATTCAGCCAGCAGC 1239
Db 388 AsnArgLeuIlyArgSerPheSerIleuAspIlySerAla----- 401
QY 1240 GGAGATCTTATCATGGCTTCTCTCATCAGAGATGTTTGAATACTCAAAACCTTCC 1299
Db 402 ----- 1300
QY 1300 ACTACTGTGATGGAGCCAC-----AAGTATGCCAGTTTC 1335
Db 406 ArgArgProAspGlyProGlyProProAspProGlyGlnAlaProIlyLeuGlnIlySer 425
QY 1336 ---TCCCT-----GTTCAAGAACTATGAGAGACAGCTCCCAACCAAGCTCTGAT 1383
Db 426 AspSerProSerGlyAlaAlaLeuGlyLeuSerSerProSerProAsp-----SerProAsp 444
QY 1384 AAGGAGAGAGCAGCATCCCAAGAGAGCTGACAGCCGACGCTTTCAGACAGCCAGCAGC 1443
Db 445 AlaAlaProGlnAlaArgProArgProArgArgArgPro----- 459

```

```

QY 1444 AAGGATTTGATTCGCTGAGAAACAGACAGTGGCACCCGACAGAGTCCCTTTATCT 1503
Db 460 -----ProAlaGlySerProAlaArg-----Ser 467
QY 1504 CCACTGATTCGAAGTGGAGCGGTGAGAGCAATTAACAACAGCTTCTTTGGGC--- 1560
Db 468 ProAlaHisSerLeuGly-----LeuAsnIleGlyAsp 478
QY 1561 -----CTTCCACAGCAGCAGACAGCACTCAAGAACTGTGTGCTGGCTTTAAGGC 1614
Db 479 AlaAlaArgGlnThrProArgHisGlyLeuSerAlaLeuSerAlaProGlyLeuProGly 498
QY 1615 -----TGCACTCGATATCTTGGCCCGCCAGACCTCT 1647
Db 499 ProGlyGlnProAlaGlyProGlyAlaIlyrAlaProProLeuAsnSerPro-----Gly 516
QY 1648 ACCCTTCCCTGACAGCAGCTGTATTTTCCACAGATCTTCAACTTACTCTGACC 1707
Db 517 ThrProSerProAspGlyProIlyProIlyCysPheSerProGln----- 529
QY 1708 TCAGCATCTACGAGAGCAGTGCATTTACTCTGCTTACAGCTGACGACGCTGACCACT 1767
Db 530 GlyAlaGlnIlyAlaGlyIlyGlyValIleuPheAlaProPheGlyArgAlaGlyAlaProGly 549
QY 1768 TGCAGACACCAAGTCTATTCTGTGCGCAGCGCAGAGACCAAGTACAGAGTACTCG 1827
Db 550 ProGlyIlyGlySerAspLeuArgArgArgGlnAlaAlaArgAlaGlnIlyProArgAspAla 569
QY 1828 CCGCGAGCTGCGCATGAGAGACCCCTTGAAGAGATTAAAGCAGAGAGCTGACCA 1887
Db 570 ArgThrGlyIlyrProGlnIlyProAlaProGlnIlyrGlnIlyrGlnIlyrGlnIlyrGlnIlyr 589
QY 1888 ATGAATTTGAGAGAGCATCATGTCAGAGAACAGTACCG-----GAAGACTGGGAAA 1944
Db 590 MetGlnPheGlnIly-----MetValGlnIlyArgAlaArgGlyGlnIlyLeuAlaAla 608
QY 1945 GTGGGAGTCACTTACCTTTTGGGCGAGATGAATCATTTGAGTCTCC 1995
Db 609 LeuGlyIlyrGlnAlaSerPheSerGlySerValGlnValIleGlnValSer 625

```

RESULT 3

DUS8_MOUSE

AC 009112; STANDARD, PRT; 663 AA.

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)

GN DUSP8 OR NTP1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=96311565; PubMed=8731137;

RA Theodosion A.M., Rodriguez N.R., Nesbitt M.A., Ambrose H.J.,

RA Paterson H., McCellan-Arnold E., Boyd Y., Leversha M.A., Owen N.,

RA Blake D.J., Ashworth A., Davies K.B.;

RT "A member of the MAP kinase phosphatase gene family in mouse

RL Hum. Mol. Genet. 5:675-684(1996).

CC -FUNCTION: This protein shows both activity toward tyrosine-protein

CC phosphate as well as with serine/threonine-protein phosphate (By

CC similarity).

CC -CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein

CC tyrosine + phosphate.

CC -CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = protein +

CC phosphate.

CC -SUBCELLULAR LOCATION: Cytoplasmic and nuclear.

CC -TISSUE SPECIFICITY: Expressed predominantly in brain and lung.

CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL: X95518; CA64772.1; -.
 DR HSSP: Q16828; 1MKP.
 DR MGD: MGI106626; Dusp8.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSRG; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00195; DSPC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS50206; RHODANESE_3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolyase; Nuclear protein.
 FT DOMAIN 23 138 RHODANESE.
 FT DOMAIN 162 432 PROTEIN-TYROSINE PHOSPHATASE.
 FT DOMAIN 452 459 POLY-ARG.
 FT DOMAIN 555 558 POLY-SER.
 FT DOMAIN 559 576 POLY-GLY.
 FT DOMAIN 577 600 POLY-SER.
 FT DOMAIN 311 552 PRO-RICH.
 FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
 SIMILARITY).
 SQ SEQUENCE 663 AA; 68847 MW; 416F429A12C1FA7C CRC64;

Alignment Scores:
 Pred. No.: 2.89e-74 Length: 663
 Score: 1302.00 Matches: 318
 Percent Similarity: 57.38% Conservative: 102
 Best Local Similarity: 43.44% Mismatches: 176
 Query Match: 36.02% Indels: 136
 Gaps: 26

US-10-029-345a-108_copy_538_2532 (1-1995) x DUS8_MOUSE (1-663)

QY 1 ATGGCCCATGATGATGGAATCAATTTGTT---ACTGAGAGTTGGTGGCTGCTGCTG 57
 Db 1 MetAlaGlyAspArgLeuProArgLysValMetAspAlaLysLysLeuAlaSerLeu 20
 QY 58 GAAAGTGAACGGAAGAAAGTGGTCTAATGATGAGCGGCATTTGGAATCAATACA 117
 Db 21 ArgGlyGlyProGlyGlyProLeuValLleAspSerLysSerPheValGlyLysSer 40
 QY 118 TCCCATATTTGGAAGCATTAATATCACTGCTCCAGGCTTAAGAACGAGTTGCA 177
 Db 41 CyHieValLeuSerSerValaenLleCySerSerLysLeuValLysArgLysGln 60
 QY 178 CAGACAAAGTGTAAATTAATACAGAGCTATCCAGCATTCAGCAAACTAAGTTGACAT 237
 Db 61 GlnGlyLysValThrLleAlaGlnLeuLleGlnProLalThrArgSerGlnValaAspAla 80
 QY 238 GATTGACGTGAGAGGTTGATTAGTATGATCAAGAGCTCCCAAGATGTGGCTGCTGCT 297
 Db 81 ThrGlnProLlnAspValLleValLysPheLlnSerThrArgAspAlaSerValLleAla 100
 QY 298 TCAGACTGTTTCTGACTGATCTTCTGGTAACTGAGAAAGAGACTCAACTGTTTAC 357
 Db 101 AlaAspSerPheLeuSerLleLeuSerLysLeuAspGlyCysPheAspSerValaLla 120

QY 358 CTGCTTGACAGAGTGGATTGCTGAGATTCTCTGTTGTTCCCTGGCCCTGTGAGAGAAA 417
 Db 121 IleLeuThrGlyLysPheAlaThrPheSerSerCysPheProGlyLeuGlyGlyLys 140
 QY 418 TCCACTGATCCCTTACC---TGCATTTCTACGCTTTGCTTACTGTTGCAATTGGG 474
 Db 141 ProLalThrLeuProSerMetSerLeuSerGlnProCysLeuProValProSerValGly 160
 QY 475 CCAACCGGAATTTCTCCCAATCTTTATCTTGGCTGCCAGGAGATGCTCCACACAGAG 534
 Db 161 LeuThrArgLleLeuProLlnSerLysLysLysLysLysLysLysLysLysLysLys 180
 QY 535 CTGATACAGAGAGATGGATTTGATTGTTAAATGACGACTATACCTGTCACAAAGCT 594
 Db 181 LeuMetThrGlnAsnGlyLleSerTyrValLeuAsnAlaSerAsnSerCysProLysPro 200
 QY 595 GACTTTATCCCGAGTCTCATTTTCTGCGTGCCTGTGAATGACAGCTTTTGTAGAAA 654
 Db 201 AspPheLleCysGlySerArgPheMetArgLleProLleAsnAspAsnLysCysGlyLys 220
 QY 655 ATTTTGGCGGTGTTGACAAATCAGTAGATTCTTATGAGAAAGCAAAAGCTCCATGGA 714
 Db 221 LeuLeuProThrLeuAspLysSerLleGlnPheLleAspLysAlaLysLeuSerSerCys 240
 QY 715 TGAATCTAGTGCATGTTAGCTGGGATCTCCGCGTCCGCAACATGCTATGCTTAC 774
 Db 241 GlnValLleValIleCysLeuAlaGlyLleSerArgSerLalThrLleAlaLleAlaTyr 260
 QY 775 ATCATGAAGAGATGAGATGATCTTTAGATGAATGAGATTGATTTGTGAAGAAAAAGA 834
 Db 261 IleMetLysThrMetGlyMetSerSerAspAspAlaLysArgPheValLysAspArg 280
 QY 835 CCTACTATATCTCCAAATCTTCAATTTTCTGGCCCAACTCTGGAATGAGAAAGATT 894
 Db 281 ProSerLleSerProAsnPheAsnPheLeuGlyGlnLeuLeuGlyLysLysSerLeu 300
 QY 895 AAGAACCAAGCTGGA-----GCATCAGGCGCCAAAGCAAAAGCTCAAGCTGCTGACCTG 948
 Db 301 LysLeuLeuAlaLalLeuGlnThrAspGlyPro-----HisLeu 313
 QY 949 GAGAGCGCAATGAACCTGTCTGCTGTCTCAGAGGTGAGACAGAAAGCAGACGCC 1008
 Db 314 GlyThrPro---GlnProLeuMetGlyProLalAlaGlyLleProLeuProArgLeuPro 332
 QY 1009 CTCAGTCCACCTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
 Db 333 ---ProSerThrSerGlnSerAlaLalThrGlySerGlnLalAlaThrAlaLalArg 350
 QY 1051 -----GGACAAAGGCCCGGTGATCCGCGCAGCGGCGCAGCGCTG 1089
 Db 351 GlnGlySerProSerAlaGlyGlyAspAlaProLle---ProSerThrAlaProLalThr 369
 QY 1090 CCCAGCGTGCAGCGTCTGCTGTTAGAGACAGCCGCTGTGATCAGGCGCTCAGTGGCTG 1149
 Db 370 SerAlaLeuGln-----GlnGlyLeuArgGlyLeu 379
 QY 1150 CACTGTCCGACAGACAGCTGGAAGACAGCATTAAGCTCAAGGTTCTTCTCTGAT 1209
 Db 380 HisLeuSerSerAspArgLeuGlnAspThrAsnArgLysLysArgSerPheSerLeuAsp 399
 QY 1210 ATCAATTCAGTTTCATATTCAGCCAGCATGCGAGCATCTTACATGCTCTCTCATCA 1269
 Db 400 IleLysSer---AlaTyrAlaProSerArgArgProAspPheProGlyProProAspPro 418
 QY 1270 GAAGATGCTTGGAAATACAAACTTCCACTACTCTGATGGAGCAACAAAGTTATGC 1329
 Db 419 GlyGlnAlaProLysLeuCysLysLeuAspSerProSerGlyGlyThrLeuGlyLeu--- 437
 QY 1330 CAGTTTCCCTGTTAGAGAACTATGAGACAGACTCCGAAACCAAGTCTGATTAAGAG 1389
 Db 438 -----ProSerProSerProAspSerPro 445
 QY 1390 GAGCCAGATCCCAAG-----AAGCTGACAGCGCGAGGCTTACAGACGCCAG 1440


```

Db 258 ArgProserIleSerProAenPhenMetCylGlnLeuGluTyrGluAenVal 277
Qy 892 ATTAAGACCAAGACTGGAGCATCAGGCGCAAGCAACTCAAGCTGGACCTTGAG 951
Db 278 Leu-----IleLysAspHisValLeuAspTyrAsn 287
Qy 952 AAGCAATGAACCTGTCCTGCTGCTCAGAGGCTGACAGAAAGCAAGCCGCTTC 1011
Db 288 GlnAlaSerArgProHisArgHisMetAspTyrTyrTyrProSerAspLeuCyPro 306
Qy 1012 AGTCCACCTGTCGCACTGCTACCTCAGAGGACAGAGCAAGAGCCGCTGATCCC 1071
Db 307 ---ProLysValProLysSerHisSerHisLancysValPheProGlySerThrHisAsp 325
Qy 1072 GCCAGCGTCGCCAGCGTCGCCAGCGTCAGCGCTGTTAGAGACAGCCGCTGTA 1131
Db 326 GluSerSerProSerSerProSerValSerLugLysSerAlaHisSerGluPro 343
Qy 1132 CAGGCGCTCAGTGGGCTGCACCTGTCGCGACAGAGGCTGGAAGACATAGCTCAG 1191
Db 343 ----- 343
Qy 1192 CGTTCTTCTCTGTGATATCAATCAGTTTCATATTCAGCAGCATGCGACATCCTTA 1251
Db 344 -----GluThrSerSerSerAlaHisSer--- 351
Qy 1252 CATGCTCTCTCTCATCAGAGATGCTTGGATATCAAAAGCTTCAGTCTGAGAT 1311
Db 352 ---SerSerThrHisSerAlaProProSerMetProSerThrHisSerGluGln 368
Qy 1312 GGGACCAACAGATGATGCGAGTTCTCCCTGTTGAGAACTATGAGAGCAGACTCCGGA 1371
Db 369 GlyThrSer-----SerGlyThrValAsnValAsnGlyLysArgAsnMet 383
Qy 1372 ACAGAGT-----CCTGATPAGAGAGAGAGCC---AGCATCCCCAAGAGAGCTGAG 1416
Db 384 ThreMetAspLeuGlyLeuProHisArgProLysAlaLeuGlyLeuProSerArgIleGly 403
Qy 1417 ACCGCC-----AGGCTTTCAGAGAGC---CAGAGCAAGCGATTCAGTTCGTCAGA 1464
Db 404 ThreSerValAlaGluLeuProSerProSerThrGluLeuSerArgLeuSerPhe-AsnG1 423
Qy 1465 ACCAGCAGCAGTGGCAGCCGCGCAGAGTCTTTATCTCAGCTGATCGAAGTGGAGC 1524
Db 423 YProGluAlaIleAlaProSerThr---ProIleLeu----- 434
Qy 1525 GTGAGAGACATPACACACAGCTTCTTTGCGCTTTCACACAGCAGCAGCAGCTTC 1584
Db 435 ---AsnPheThrAsnProCyPheAsnAsnSerProIleIleProValAlaSerSerSe 452
Qy 1585 ACAGAGTCTGCTGCGCTGGGCTTAAAGGCTGGGACTCGGATATCTTGGCCCCCAGACC 1644
Db 452 ValGluValIle----- 456
Qy 1645 TCATACCTTTCCTGACAGCAGCTGATATTTGCCACAGAGTCTTCACACTTCTACTCT 1704
Db 457 -LeuThrLeuProThrProAlaIa-----SerSerSerSerSerThrSe 471
Qy 1705 GCGTCAGACCA-----TTCAGGAGAGAGAGCGCAGT 1734
Db 471 rSerGluProSerPheAspPheSerSerPheGluSerSerSerSerSerIleValIa 491.
Qy 1735 TACTGTGCTTACAGCTGCGACAGCTGCCACTTGGCGAGACCAAGTCTTATTCGTGCGC 1794
Db 491 IgluAsnProPhePheAlaSerThrGluValProAlaGlySerSerSerIleSerThrPr 511
Qy 1795 AGGCGGAGAGAGCAAGTGAATGACAGAGCTGACTCGGCGGAGCTGAGTGAAGAGAGCCC 1854
Db 511 oSerGly---SerGlnSerThrProAlaSerAlaSerSerSerAlaHisSerArgCyAr 530
Qy 1855 TTGAAA 1861
:::

```

```

Db 530 GmetLys 532
RESULT 5
ID DUS6_HUMAN STANDARD; PRT; 361 AA.
AC Q16828; 075109; Q9BSH6;
DT 01-NOV-1997 (Ref. 35, Created)
DT 01-NOV-1997 (Ref. 35, Last sequence update)
DT 10-OCT-2003 (Ref. 42, Last annotation update)
DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
DE phosphatase 3) (MKP-3) (Dual specificity protein phosphatase PYST1).
GN DUSP6 OR MKP3 OR PYST1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Forebrain;
RX MEDLINE=96312959; PubMed=8670865;
RA Groom L.A., Sneddon A.A., Aleasi D.R., Dowd S., Keyse S.M.;
RT "Differential regulation of the MAP, SAP and RK/p38 kinases by Pyel1,
RT a novel cytosolic dual-specificity phosphatase.";
RL EMBO J. 15:3621-3632(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Liver;
RX MEDLINE=99077745; PubMed=9858808;
RA Furukawa T., Yatsuno K.T., Yousef E.M., Abe T., Yokoyama T.,
RA Fukuhige S., Soeda E., Hoshi M., Hayashi Y., Sunamura M., Kobari M.,
RA Horii A.;
RT "Genomic analysis of DUSP6, a dual specificity MAP kinase phosphatase,
RT in pancreatic cancer.";
RL Cytogenet. Cell Genet. 82:156-159(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT VAL-114.
RC TISSUE=Colon, Kidney, Skin, and Stomach;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grioux L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Diatchenko L., Marziani K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Udell T.B., Toshyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalka U., Smilins D.E.,
RA Schmeckel A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 204-347.
RX MEDLINE=99140299; PubMed=10048930;
RA Stewart A.E., Dowd S., Keyse S.M., McDonald N.O.;
RT "Crystal structure of the MAPK phosphatase Pyel1 catalytic domain and
RT implications for regulated activation.";
RL Nat. Struct. Biol. 6:174-181(1999).
CC -!- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
CC family.
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

```



```

Db 285 LysAsnGlyValLeuValHisCysLeuAlaGlyLeuSerArgSerValThrValThr 304
Qy 766 ATGCGCTACATCATGAAGAGATGACATCTTTAGATGACCTTACAGATTGTGAAA 825
Db 305 ValAlaTyrLeuMetGlnTylsLeuAsnLeuSerMetCysAspAlaTyrAspIleValTyls 324
Qy 826 GAAAAAGACCTACTATCTATCTCCAACTTCTGTGGCCAACTCTGACACTATGAG 885
Db 325 MetTylsYserAsnIleSerProAsnPhaAsnMetGlyGlnLeuAsnAspPheGlu 344
Qy 886 AAGAAGATTAGAACAGACTGAGACGACGCA 921
Db 345 ArgThrLeu-----GlyLeuSerSerPro 352

RESULT 6
DUSA_HUMAN STANDARD; PRT; 482 AA.
AC G9Y6M6;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 5) (MAP kinase
DE phosphatase 5) (MKP-5).
DE DUSP10 OR MKP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=99321929; PubMed=10391943;
RA Tanoue T., Moriguchi T., Nishida E.;
RT "Molecular cloning and characterization of a novel dual specificity
RT phosphatase, MKP-5."
RL J. Biol. Chem. 274:19949-19956(1999).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=2065165; PubMed=10597297;
RA Theodosiou A., Smith A., Gillieron C., Arkinstall S., Ashworth A.;
RT "MKP5, a new member of the MAP kinase phosphatase family, which
RT selectively dephosphorylates stress-activated kinases."
RL Oncogene 18:6981-6988(1999).
RN 13
RP SEQUENCE FROM N.A.
RT TISSUE=Brain, Lung, and Testis;
RC MEDLINE=22388257; PubMed=12477932;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toehlyuk S., Carantini P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.U., Abrahamson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fehay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywiński M.I., Skalska U., Smalins D.E.,
RA Schenker A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Involved in the inactivation of MAP kinases. Has a
CC specificity for the MAPK1/MAPK2/MAPK3/MAPK4 subfamily.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +

```

```

CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AB026436, BA81668.1, -.
DR EMBL, AF179212, AAD51857.1, -.
DR EMBL, BC031405, AAH31405.1, -.
DR HSSP, Q16828, MKP.
DR Genew, HGNC:3065, DUSP10.
DR GO, GO:0005737, Cytoplasm, TAS.
DR GO, GO:0005634, Cytosol, TAS.
DR GO, GO:0004721, F-protein phosphatase activity, TAS.
DR GO, GO:0007254, P-JNK cascade, TAS.
DR GO, GO:0006470, P-protein amino acid dephosphorylation, TAS.
DR GO, GO:0006950, P-response to stress, TAS.
DR InterPro, IPR000340, DS_phosphatase.
DR InterPro, IPR008343, MAP_kinase.
DR InterPro, IPR001763, Rhodanese-like.
DR InterPro, IPR000387, Tyr_phosphatase.
DR Pfam, PF00782, DSPC, 1.
DR Pfam, PF00581, Rhodanese, 1.
DR PRINTS, PR01764, MARKPHPTASE.
DR SMART, SM00195, DSPC, 1.
DR SMART, SM00450, RHOD, 1.
DR PROSITE, PS02066, RHODANASE_3, 1.
DR PROSITE, PS00383, TYR_PHOSPHATASE_1, 1.
DR PROSITE, PS00056, TYR_PHOSPHATASE_2, 1.
DR PROSITE, PS00054, TYR_PHOSPHATASE_DUAL, 1.
DR KMW, Hydroxylase; Nuclear protein.
DR FT, DOMAIN 168 285 RHODANASE.
DR FT, DOMAIN 384 453 PROTEIN-TYROSINE PHOSPHATASE.
DR FT, ACT_SITE 408 408 PHOSPHOCYSTEINE INTERMEDIATE (BY
DR FT, ACT_SITE 408 408 SIMILARITY).
DR SO, SEQUENCE 482 AA; 52642 MW; A8CB74ABF9498CD4 CRC64;

Alignment Scores:
Pred. No.: 4,56e-22 Length: 482
Score: 469.00 Matches: 107
Percent Similarity: 57.72% Conservative: 65
Beet Local Similarity: 35.91% Mismatches: 96
Query Match: 12.97% Indels: 30
DB: 1 Gaps: 7

US-10-029-345A-108_COPY_538_2532 (1-1995) x DUSA_HUMAN (1-482)
Qy 79 CCGCTATGATGATGACCGGCGCATTTGTGGAATACATATCCACATTTTGAAGCCATT 138
Db 173 ValIleIleAspCysArgProPheMetGlnTyrAsnLysSerIleIleGlnTylValAl 192
Qy 139 AATATCACTGTCGCC---AAGCTTGAAGCGAAGTTCACACAGACAAGAGTTAAAT 195
Db 193 HisIleAsnCysAlaAspTylsLeuSerArgValGlnGlnIleTylValIleThrVal 212
Qy 196 ACAGACTCATCCAG---CATTCAGCGAACAATTAAGTTGACATTTGACACTCAGAG 252
Db 213 LeuAspLeuIleSerCysArgGlnTylsValAspSerPheTylArgTylLeuPheSerTylGln 232
Qy 253 GTTGTGATTAGATTAAGCTCCCAAGATGTGGCTCTCTCTTCAACACTTTTCTC 312
Db 233 IleIleValTylAspIleAsnThrArgGlnTylProSerArgValLeuProSerGlnProLeu 252
Qy 313 ACTGACTCTGGGTAACTGGAAGAGAGACTTCAACTGTGTCACCTGCTTGACAGGTGG 372

```


QY 79 CTGCTAATGATAGCCGCGCATTTGTGATATACATATCCAGCATTTTGAAGCAT 138
 Db 174 ValIleIleAspCybargProPheMetGluTyrAsnYserHisIleGlnGlyAlaVal 193
 QY 139 AATATCAACTGCTCC--AGCTTATGACGGAAGGTTGACAGCAAGATGTTAAT 195
 Db 194 HsIleIleAsnCybAlaAspGlySileSerArgArgArgLeuGlnGlnGlySileThrVal 213
 QY 196 ACAGAGCTCATCCAG--CATTCAGCAACATTAAGGTTGACATTCAGTCCAGAG 252
 Db 214 LeuAspLeuIleSerCybargGluGlyIlyAspSerPheYllyAspGlyLeu 253
 QY 253 GTGTGATTTAGCATCAACAGCTCCAGATGTTGCTCTCTTCAGCATGTTTCTC 312
 Db 234 IleIleValTyrAspGlnAsnThrAsnGlnProSerArgValThrProSerGlnProLeu 253
 QY 313 ACTGATCTTGGGATTAATGAGAGAGAGCTTCAACTGTTGACAGAGTGGG 372
 Db 254 HsIleIleValLeuIleuSerLeuIlyAspGluGlyIlyLeuIleuValLeuIlySglY 273
 QY 373 TTGTCGAGTCTCTCTGTTTCCCTGCGCTCTGTGAGAGAAA----- 417
 Db 274 LeuSerSerPheYllySglAsnHisIleGlyAsnLeuCybAspAsnSerLeuGlnGln 293
 QY 418 -----TTCACCTGAGCCCTTACCTGATTTCT 444
 Db 294 CybArgGluValGlyIlyGlyAlaSerAlaIleSerSerMetLeuProGln-----Ser 311
 QY 445 CAGCCTTGTCTTACT--GTTGCCAATTTGGGCAACCCGAATTTTCCCAATTTTAT 501
 Db 312 ValProThrThrProAspIleGluAsnAlaGluIleThrProIleuProPheLeuPhe 331
 QY 502 CTGGGCTCCAGAGAGATGTCCTCCACAGAGAGCTATACAGAGATGAGATTTGTTAT 561
 Db 332 LeuGlyAsnGlnIleAspAlaGlnAspLeuAspThrMetGlnArgLeuAsnIleGlyTyr 351
 QY 562 GTCTTAATGCCAGC-----TATACCTGTCGAAGCTGATTTATCCCC 606
 Db 352 ValIleAsnValThrThrIleuProLeuTyrHisIlyGlyIlySglYLeuPhe----- 369
 QY 607 GAGTCTTATTCCTGCGTGTGCTGTCGATAGAGAGCTTTTGGAGAAATTTTCCGCTG 666
 Db 370 -----AsnTyrIlyAspGluProAlaThrAspSerAsnIlySglAsnLeuArgTyr 387
 QY 667 TTGGACAAATCAGTATTCATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726
 Db 388 PheGlnGluAlaPheGluPheIleGlnGluAlaHisIleGlnCybGlySglYLeuLeuIle 407
 QY 727 CACTGTTTGGTGGAGATGCTGCGCTGCGCAGCAGATGCTTATCCCTACATCAGAGAGAG 786
 Db 408 HsIleGlnIleAlaGlyAlaSerArgSerAlaThrIleValIleAlaTyrLeuMetIlySHis 427
 QY 787 ATGACATGCTTTAGATAGAGCTTACAGATTTGTGAGAGAGAGAGAGAGAGAGAGAG 846
 Db 428 ThrArgMetThrMetThrAspAlaTyrIlyPheValIlySglYAspArgProIleIleSer 447
 QY 847 CCAACTTCAATTTTCTGGGCGCAACTCTGAGCTATGAGAGAGAGAGAGAGAGAGAG 900
 Db 448 ProAsnLeuAsnPheMetGlyIleuLeuGluPheGluIlyuAspLeuAsn 465

RESULT 8

DUS6_MOUSE STANDARD; PRT; 381 AA.
 AC Q9DBB1; Q9D7L4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
 DE phosphatase 3) (MKP-3).
 OS DUSP6 OR MKP3.
 GN Mus musculus [Mouse].
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver, and Tongue;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Kono H., Balderelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bulc C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,
 RA Myszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=23388257; PubMed=1247932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Zeeberg B., Wagner L., Shenmen C.F., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.M., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Maira M.A.;
 RT Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
 CC family (By similarity).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL, restriction -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-bb.ch).
 CC -----
 CC EMBL; AK005062; BAB23786.1; -;
 CC EMBL; AK009131; BAB26093.1; -;
 CC EMBL; BC003869; AAH03869.1; -;
 CC HSSP; Q16828; IMKP.
 DR MGD; MGI:1914853; Dusp6.

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).

DR EMBL; X94185; CA63895.1; -
 DR EMBL; U46227; AB06202.1; -
 DR HSSP; Q16828; IMKP.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANESE_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydroxylase.
 FT DOMAIN 30 148 RHODANESE.
 FT DOMAIN 206 381 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 293 293 PHOSPHOCYSTEINE INTERMEDIATE (BY
 SIMILARITY).
 SQ SEQUENCE 381 AA; 42318 MW; C511B0CB6F2868 CRC64;

Alignment Scores:

Pred. No.: 7.26e-22 Length: 381
 Score: 465.50 Matches: 112
 Percent Similarity: 53.01% Conservative: 64
 Best Local Similarity: 53.73% Mismatches: 107
 Query Match: 12.88% Indels: 49
 DB: Gaps: 8

US-10-029-345a-108_COPY_538_2532 (1-1995) x DUSE6_RAT (1-381)

QY 55 CTGGAAGTGGAAAGCGAAAAAGTCTGTAATGATGCGCGCATTTGTAATCAAT 114
 DB 27 LGGGLYUENGLYVANGIUGLNULEULEMEARPCYARPGROGNGIUGLEUYGLU 46
 QY 115 ACATCCCAATTTTGGAAAGCATTAATATCAACGCTCCCAAGCTTGAAGAGGTTG 174
 DB 47 SGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 66
 QY 175 CAACAGGACAAAGTGAATTAATTAAGAGCTATCCAGCAT--TCAGCGAAACATTAAGTT 221
 DB 67 GlnlyseLYuEnleuProValArgAlaLeuPheThArgCyseLYuEnleuArgAlaArgPhe 86
 QY 232 GACATTGATTGC--AGTCAAGAGTGTGATGATTAAGATCAAGATCCCAAGAT--GTT 285
 DB 87 ThArgArgCyseLYuEnleuThArgThValValLeuLYuEnleuArgAlaArgTArgAla 106
 QY 286 GCGCTCTCTCTTCAAGAGCTTTTTCAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCT 345
 DB 107 GlnAenThArgLYuEnleuSerValLeuLYuEnleuLYuEnleuLYuEnleuArgLYu 126
 QY 346 AACTCTTCTCAAGTCTGAGAGGAGTGTGATGATTAAGATCAAGATCCCAAGATCC 405
 DB 127 CyArgAlaTherLYuEnleuLYuEnleuLYuEnleuLYuEnleuLYuEnleuLYuEnleu 146
 QY 406 TGTGAAGAAATGCACTTGAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 462
 DB 147 CyseLYuEnleuThArgThValValLeuLYuEnleuLYuEnleuLYuEnleuLYuEnleu 164
 QY 463 GCGCAACATTTGGG----- 474
 DB 165 LeuGLYUENGLYGLYLeuArgLYuEnleuSerSerSerSerSerSerSerAlaLYuEnleu 184
 QY 474 ----- 474
 DB 185 AspArgAspProAsnSerAlaThrArgSerArgGLYSerProLeuSerAsnSerGlnPro 204

QY 475 -----CNAACCGGATCTTCCCATTTATCTTGGCTGCGAGGAGATGTCCTCAAC 528
 DB 205 SerPheProValGlnLYuEnleuProPheLeuLYuEnleuLYuEnleuLYuEnleuLYuEnleu 224
 QY 529 AAGAGCTGATACAGAGAGATGGATTTGATGTATGTTAAATCCAGCTTATCTGCA 588
 DB 225 LeuArgValLeuGlnLYuEnleuLYuEnleuLYuEnleuLYuEnleuLYuEnleuLYuEnleu 244
 QY 589 AAGCTT--GACTTATCCCGAGTCTCATTTCTCGCGTGGCTGGAATGACAGCTTT 645
 DB 245 AsnLeuPheGlnAsnAlaGlyGlnPheLYuEnleuLYuEnleuLYuEnleuLYuEnleuLYuEnleu 264
 QY 646 TGTGAAGAAATTTGGCGGCTTGGCAAAATCGTATGATTTCAATGAAAGAAAGCC 705
 DB 265 SerGlnAsnLeuSerGlnPheProGlnAlaLYuEnleuSerPheLeuArgAlaArgLYu 284
 QY 706 TCCATGATGATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 765
 DB 285 LysArgCyseLYuEnleuValHisCyseLYuEnleuLYuEnleuLYuEnleuLYuEnleuLYuEnleu 304
 QY 766 ATCGCTTACATCATGAAGAGATGACATGCTTTAGATGAGCTTACGATTTGAAA 825
 DB 305 ValAlaLYuEnleuMetGlnLYuEnleuAsnLeuSerMetAsnArgAlaLYuEnleuLYuEnleu 324
 QY 826 GAAAAAGACCTTATCTTCCAACTTCAATTTCTGGCGCACTCCGCACTATGAG 885
 DB 325 MetLYuEnleuSerAsnLYuEnleuSerProAsnPheAsnPheMetGlnLYuEnleuAspPheGln 344
 QY 886 AAGAGATTAAGAACCACTGAGCATGAGGCGCA 921
 DB 345 ArgThLeu-----GlyLeuSerSerPro 352

RESULT 10
 ID DUS4_HUMAN STANDARD; PRT; 394 AA.
 AC 01315; Q13524;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
 phosphatase-2) (MKP-2) (Dual specificity protein phosphatase hVH2).
 GN DUSP4 OR MKP2 OR VH2.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95221370; PubMed=7535768;
 RA Guan K.-L., Butch E.,
 RT "Isolation and characterization of a novel dual specific phosphatase,
 RT hVH2, which selectively dephosphorylates the mitogen-activated
 RT protein kinase."
 RT J. Biol. Chem. 270:7197-7203 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96198119; PubMed=8626452;
 RA Chu Y., Solaki P.A., Khosravi-Far R., Der C.J., Kelly K.,
 RT "The mitogen-activated protein kinase phosphatases PAc1, MKP-1, and
 RT MKP-2 have unique substrate specificities and reduced activity in vivo
 RT toward the ERK2 sevenmaker mutation."
 RT J. Biol. Chem. 271:6497-6501 (1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC Tissue=skin, and Uterus;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heifetz F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bernaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mallya S.J.,
RA Boesak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
RA Villalon D.K., Wozny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy U., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.W.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases
CC ERK1 and ERK2.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, U21108; AA85119.1; -;
DR EMBL, U48807; AAC50452.1; -;
DR EMBL, BC002671; AA02671.1; -;
DR EMBL, BC014565; AA014565.1; -;
DR HSSP, Q16828; IMKP.
DR GeneW, HGNC:3070; DUSP4.
DR MIM, 602747; -;
DR GO, GO:0005634; C:nucleus; TAS.
DR GO, GO:0008330; F:protein tyrosine/threonine phosphatase acti. . .; TAS.
DR GO, GO:0000165; P:MAPKK cascade; TAS.
DR GO, GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR InterPro, IPR000340; DS_phosphatase.
DR InterPro, IPR008343; MAPK_phosph.
DR InterPro, IPR001763; Rhodanese-like.
DR Pfam, PF00782; DSPc; 1.
DR Pfam, PF00581; Rhodanese; 1.
DR PRINTS, PR01764; MAPKPHPTASE.
DR SMART, SM00195; DSPc; 1.
DR SMART, SM00450; RHOD; 1.
DR PROSITE, PS00206; RHODANSE 3; 1.
DR PROSITE, PS00383; TYR_PHOSPHATASE 1; 1.
DR PROSITE, PS0056; TYR_PHOSPHATASE 2; 1.
DR PROSITE, PS0054; TYR_PHOSPHATASE_DUAL; 1.
DR HydroLase, Nuclear protein.
FT DOMAIN 41 159
FT ACT_SITE 197 394
FT ACT_SITE 280 280
FT CONFLICT 111 111
SQ SEQUENCE 394 AA; 42953 MW; 060397175986952E CRC64;
Alignment Scores:
Pred. No.: 1,06e-20
Score: 447.00
Percent Similarity: 50.38%
Best Local Similarity: 30.89%
Query Match: 12.37%
Length: 394
Matches: 122
Conservative: 77
Mismatch: 141
Indels: 55

DB: 1 Gaps: 14
US-10-029-345a-108_copy_538_2532 (1-1995) x DUS4_HUMAN (1-394)
QY 45 GGTGGCTGCTGGAAGATGGAACGGA-----AAAGT 77
DB 26 G1G1G1A1A1G1G1G1-SerG1YSerH1sg1YThrLeuG1YLeuProSerG1Yg1Yg1Scy 45
QY 78 GCTGCTAAATGATAGCCGCGCATTTGTGGAATACATATCATCCCAATTTTGAACCCAT 137
DB 45 SLeuLeuLeuSPrCySargProPheLeuAlaHisSerAlaG1YThrLeuLeuSerVa 65
QY 138 TAAATCAACTGCTCCAAAGCTTATGAGACGACAGCTTCAACAGACAAAGTTAATAC 197
DB 65 IAsnValArgCySasnthrIleValAlaArgaGarg--AlaYgSerValSerLeuG1 84
QY 198 AGAGCTCATC-----CAGCATTCAGCGAACAATAAGTTGACATTTGACATTCAGACAGA 251
DB 84 uGlnIleLeuProAlaG1G1G1G1uValAlaArgAlaArgLeuArgSerG1YLeuYrSerAl 104
QY 252 GGTGTAGTTTACAGATCAAGATCCCAAGATGTGCTCTCTTCAACTGTTTCT 311
DB 104 aValIleValAlaYrAspIuArgSerProArgAlaG1uSerLeuArgG1uAspSerThrVa 124
QY 312 CACTGTACTTCTGGTAAACTGAGAGAGACTTC-----AACTGTTCACCTGCTTGC 365
DB 124 IserLeuValValGlnAlaLeuAlaArgAsnAlaG1uArgThrAspIleCySLeuLeu 144
QY 366 AGGTGGATTGTGCTGATGCTCTCGTTTCCCTGAGCTTGAAGAAATGCACCTC 425
DB 144 sg1Yg1YrG1uAluArgPheSerSerG1uYrProG1uPheCySerIYrThrYsAlaLe 164
QY 426 A-----GTCCCTACCTGCATTTCTCAGCT----- 450
DB 164 uAlaAlaIleProProProValProProSerAlaThrGluProLeuAspLeuG1YCySe 184
QY 451 -----TGC--TTACCTGTGGCAACAT--GGGCCAACCCGAATTTCTCCAAATGTTTA 500
DB 184 rSerCySg1YrHrProLeuHisAspG1ng1Yg1YrProValG1uIleuProPheLeuY 204
QY 501 TCTTGGCTGCCAGGAGATGTCTCAACAGAGCTGATACAGACATGGAGTTGGTTGA 560
DB 204 rLeuG1YSerAlaYrH1sAlaAlaAlaArgaAspMetLeuAspAlaLeuG1YIleThrAl 224
QY 561 TGTGTTAAATCCGACATACCTGTCACCAAGCTGATTTATCCCGAGTCTCATTTCT 620
DB 224 aLeuLeuAsnValSerSerAspCySProAsn--HisPheG1uG1YH1eYrG1uYrIly 243
QY 621 GCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTGGCGGTGGAGCAATATCOT 680
DB 243 sCySAlleProValG1uAspAsnHisIlySAlaAspIleSerSerTrpPheMetGluAlaI 263
QY 681 AGATTTCATTGAGAAAGCAAAAGCCCTCCCAATGATGTGTGTTAGTGCATGTTTAACTGG 740
DB 263 eGluYrIleAspAlaValAlaYrAspCySargIYrG1uValIleValHisCySg1uAlaG1 283
QY 741 GATTCCTCCGCTCGCCACCATGCTATCGCTATCATCATCAAGAGATGAGCATGCTTT 800
DB 283 YIleSerArgSerAlaThrIleCySLeuAlaIYrLeuMetIleYrSAlaYrAlaG1e 303
QY 801 AGATGAAGCTTACAGATTGTGAGAAAGAAAGAAAGCTTATATCTCAAACTTCAATT 860
DB 303 uG1uG1uAlaPheG1uPheValYsg1uArgYrSerIleIleSerProAsnPheSerPh 323
QY 861 TCTGGGCCACTCTCTGACATATGAGAGAAAGATTTAGAACACAGCTGAGACATCAGGCC 920
DB 323 eMetG1Yg1uLeuLeuG1uPheG1uSerG1uValIleValIleThrSerCySAlaG1uAl 343
QY 921 AAAAGCAAACTCAAGCTGCTGACCTGAGACCCAAATGAATCTGTCTGCTGCTC 980
DB 343 aAlaSer-----ProSerG1YProLeu----- 350
QY 981 AGAGGTGAGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1040

Db 351 -----ArgGluArgGlyThrProAlaThrPro-----Thrse 362
QY 1041 AGAGGACGACGACGAAAGCCCGTGCATCCCGACGCTCCAGCGTCCGACGCTGCA 1100
Db 362 rgInPheValPheSerPheProVal---SerValGlyValHisSerAlaProSerSerLe 381
QY 1101 GCGGTGCTGTTAGAGACGCGCGCTGTCAGGCGCTCACT 1143
Db 381 uProTyLeu-----HisSerProIleThrThrSerProSer 393
RESULT 11
DUS4_CHICK STANDARD; PRT; 375 AA.
ID Q9PW71;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
phosphatase-2) (MKP-2).
GN DUSP4 OR MKP2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=20379359; PubMed=10918612;
RA Fu S.-L., Waha A., Vogt P.K.;
RT Identification and characterization of genes upregulated in cells
transformed by v-Jun.";
RL Oncogene 19:3537-3545(2000).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
dephosphorylating both Thr and Tyr residues on MAP kinases ERK1
and ERK2 (by similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL, A167296; AAD46556.1; -.
DR HSSP, Q16828; IMKP.
DR InterPro, IPR000340; DS phosphatase.
DR InterPro, IPR008343; MAPK phosph.
DR InterPro, IPR001763; Rhodanese-like.
DR InterPro, IPR000387; Tyr_phosphatase.
DR Pfam, PF00782; DSPc; 1.
DR Pfam, PF00581; Rhodanese; 1.
DR PRINTS, PR01764; MAPKPHPTASE.
DR SMART, SM00195; DSPC; 1.
DR SMART, SM00450; RHOD; 1.
DR PROSITE, PSS0206; RHODANES_3; 1.
DR PROSITE, PSS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE, PSS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE, PSS0054; TYR_PHOSPHATASE_DUAL; 1.
DR HydroLase; Nuclear protein.
KM DOMAIN 25 143 RHODANES.
FT DOMAIN 178 375 PROTEIN-TYROSINE PHOSPHATASE.

FT ACT_SITE 261 261 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SEQUENCE 375 AA; 41052 MW; 17929DDCC2BEEF1 CRC64;
SIMILARITY)
Alignment Scores:
Pred. No.: 1, 61e-20 Length: 375
Score: 444.00 Matches: 123
Percent Similarity: 49.49% Conservative: 73
Best Local Similarity: 31.06% Mismatches: 152
Query Match: 12.28% Gaps: 48
DB: 1 Gaps: 13
US-10-029-345a-108_COPY_538_2532 (1-1995) x DUS4_CHICK (1-375)
QY 10 GAGATGATTGGAATCTGAATTGTTACTGAGAGTTGCTGCTGCGAAGTGAACG 69
Db 9 GluMetGluGlySerLeu-----ArgArgLeuValGlyArgGluGluAlaSerGly 26
QY 70 GAAAAGTGTGCTATTGATTGATGCGGCGCATTTGTGGAATACATTCACATTTTG 129
Db 27 GlyArgCysLeuLeuLeuSerCysArgProPheLeuAlaHisSerAlaGlyHisLeuArg 46
QY 130 GAAGCATTAATATCAACTGCTCCAGCTTATGAGCGAAGTTGCAACGACGAAAGTG 189
Db 47 GlyAlaLeuAsnValArgCysAsnThrIleValArgArg---AlaLysGlyAlaVal 65
QY 190 TTAATTACAGAGCTATC-----CAGCATTCAGCGAAACATAGGTTGACATTGTC 243
Db 66 SerLeuGluGlnIleLeuProAlaGluGlyGluValArgAlaArgLeuArgAlaGlyLeu 85
QY 244 AGTCAGAAAGTTGATTTACATCAATCAAGCTCCAGATGTTGCTCTCTTCAGAC 303
Db 86 TyrThrAlaValAlaLeuValArgGlyArgSerProArgAlaGluAlaLeuArgAsp 105
QY 304 TGTTCCTCACTGACTTCTGAGTGAATCTGAGAAAGAGCTTC-----AACTCTGTTAC 357
Db 106 SerThrValAlaLeuValAlaLeuArgAlaLeuArgArgAspMetAlaArgAlaLeuArg 125
QY 358 CTGCTTGACAGGTGGGTTTGCTGAGTTCTCGTTGTTCCCTGCGCTCTGTGAAGAAA 417
Db 126 LeuLeuAlaGlyGlyArgGluArgPheAlaSerGluTyrProGluPheCysAlaTyrThr 145
QY 418 TCCACTGATGCTCCATCCGATTCAGGCT----- 450
Db 146 LysThrLeu-----SerSerIleSerProProSerSerAlaGluSerLeuAspLeuGly 163
QY 451 -----TGC---TTACTGTGGCCAACTT---GGGCCAAACCGAATTTCTTCCCAAT 495
Db 164 PheSerSerCysGlyThrProLeuHisAspGlnGlyGlyProValGluIleLeuProPhe 183
QY 496 GTTATCTTGAGTGCACGAGATGTCTCAACAGAGAGCTGATACAGAGAAATGGAGATT 555
Db 184 LeuTyrLeuGlySerAlaTyrHisAlaAlaArgArgAspMetLeuAspAlaLeuGlyIle 203
QY 556 GGTATGTTGTTAAATCCGACGTATACCTGTCGCAAAAGCTGACTTTATCCCGAGTCTCAT 615
Db 204 ThrAlaLeuLeuAsnValSerSerAspCysProAsn---HisPheGluGlyHisTyrGln 222
QY 616 TTCCTCGTGTGCTGCTGATGATGACAGCTTTGTGAGAAATTTTGCGGTGTGACAAA 675
Db 223 TyrLysCysIleProValGluAspAsnHisLysValAlaAspIleSerSerThrPheMetGlu 242
QY 676 TCAAGTAAATTCATTGAGAAAGAAAGCTCCATATGATGTGTTTACTGCACTGTTTA 735
Db 243 AlaIleGluTyrIleAspSerValLysGlyCysGlyArgValAlaLeuValHisCysGln 262
QY 736 GCTGGAGATCCCGCTCGGCGCAACATCGTATCGCTACATGAAAGAGATGACATG 795
Db 263 AlaGlyIleSerArgSerAlaThrIleCysLeuAlaTyrLeuMetCysLysValArgVal 282
QY 796 TCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTACTATATCTCCAACTTC 855
Db 283 LysLeuGluValAlaPheGluPheValLysGlnArgArgSerIleIleSerProAsnPhe 302


```

QY 643 TTTTGTGAGAAAATTTTCCGCTGTTGAGCAAAATCAGTACATTTTCATTGAGAAAGCAAAA 702
Db 252 HtlylalaaplleSerSerTrpPheMetGluAlaIleGluTyrIleAapAlaValLys 271
QY 703 GCCTCCAAATGATGTTGTAGTACATCTGTTAGTCTGGATCTCCGCTCCGACCATC 762
Db 272 AapCyAargllylAargValIleValHisCyGlnAlaIleGlyIleSerAgserrAlaThrIle 291
QY 763 GCTATGCGCTTACATCATGAAAGAGAGATGCTTTAGATGAAGCTTACAGATTGTG 822
Db 292 CyLeuAlaTyrIleuMetIleTylAargValAArgleuGluAlaIleAapGluPheVal 311
QY 823 AAAGAAAAGAACTTACTATATCTCCAAATTTCTGCGCCCAACTCCGAGACTAT 882
Db 312 LysGlnIArgIserIleIleSerProAphSerMetGlyGlnIleuGlnPhe 331
QY 883 GAGAAGAGATTAAAGACAGATGAGATCAGGCGCAAGAGCAAACTCAAGCTGCTG 942
Db 332 GluSerGlnValIleuThrThrSerCyAlaAlaIleGluAlaIleSer 346
QY 943 CACCTGAGAGACCAATGAATGACCTGCTGCTGCTCAGAGGAGTGACAGAAAGCGAG 1002
Db 347 -----ProSerGlyPro-----LeuArgGluAArgIlySalAThrPro 359
QY 1003 AGCGCCCTCAGTCCACCTGTGCGGACTGCTGTAACCTCAAGAGCAGAGCAAAAGCCC 1062
Db 360 ThrPro-----ThiSerIlePheValIlePheSerPhePro 370
QY 1063 GTGCATCCCGCAGCGTCCGAGCGTCCGAGCGTCCGAGCGTCCGAGCGTCCGAGCG 1122
Db 371 Val-----SerValGlyValHisAlaAlaIleProSerAhnLeuProTyrIleu-----HisSer 387
QY 1123 CGCGTGTGACAGCGCTCACT 1143
Db 388 ProIleThrThrSerProSer 394

RESULT 13
DUST_MOUSE STANDARD; PRT; 367 AA.
ID DUST1_MOUSE
AC P28563;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
DE 3CH134) (Protein-tyrosine phosphatase ERP).
GN DUSP1 OR PTPN10 OR MKP1 OR 3CH134 OR PTPN16.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=92158357; Pubmed=1741163;
RA Charles C.H., Adler A.S., Lau L.F.;
RT "CDNA sequence of a growth factor-inducible immediate early gene and
RT characterization of its encoded protein.";
RL Oncogene 7:187-190(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93360956; Pubmed=8355678;
RA Noguichi T., Metz R., Chen L., Mattei M.-G., Carrasco D., Bravo R.;
RT "Structure, mapping, and expression of erp, a growth factor-inducible
RT gene encoding a nontransmembrane protein tyrosine phosphatase, and
RT effect of ERP on cell growth.";
RL Mol. Cell. Biol. 13:5195-5205(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary gland;
RX MEDLINE=2308257; Pubmed=12477932;

```

```

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schlier G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helen F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,
RA Raba S.S., Lomellano N.A., Peters G.J., Abramson R.D., Millaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fehey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Ratterfield Y.S.N., Krzywnski M.I., Skalska U., Smilins D.E.,
RA Schenker A., Schein J.E., Jones S.J.M., Maria M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=94037096; Pubmed=8221886;
RA Sun H., Charles C.H., Lau L.F., Tonks N.K.;
RT "MKP-1 (3CH134), an immediate early gene product, is a dual
RT specificity phosphatase that dephosphorylates MAP kinase in vivo.";
RL Cell 75:1487-1493(1993).
CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
CC kinase ERK2 on both Thr-183 and Tyr-185.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- INDUCTION: By growth factors.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-eb.ch/announce/
CC or send an email to license@isb-eb.ch).
CC
DR EMBL: X61940; CAA43944.1; -
DR EMBL: S64851; AAB27882.1; -
DR EMBL: BC006967; AAH06967.1; -
DR PIR: A54681; S24411.
DR HSSP: Q16828; MKP.
DR MGD: MGI:105120; Dusp1.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; Tyr_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS: PR01764; MAPKPHPTASE.
DR SMART: SM00450; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANES_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydrolyase; Cell cycle.
FT DOMAIN 20 137 RHODANES.
FT ACT_SITE 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE.
FT MUTAGEN 258 258 C->S: LOSS OF ACTIVITY.
SQ SEQUENCE 367 AA; 39369 MW; 50B5F90FEBBD19AB CRC64;

```

Alignment Scores:


```

Pred. No.: 7,87e-20 Length: 367
Score: 433.00 Matches: 115
Percent Similarity: 46.43% Conservative: 54
Best Local Similarity: 31.59% Mismatches: 151
Query Match: 11.98% Indels: 44
DB: 1 Gaps: 8

US-10-029-345a-108_copy_538_2532 (1-1995) x DUS1_MOUSE (1-367)

QY 43 TTGGTGGCTCTGCTGAAAAGTGAACGAAAAAGTCTGTAATGATAGCCGCAATT 102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 LeuAlaGAlaLeuLeuArgLgLuAlaAlaGlnCysLeuLeuLeuAspCysArgSerPhe 32
QY 103 GTGGAATACATATACATCCCACTTTTGGAGCAATTAATATATCACTGCTCAAGCTTATG 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 PheAlaPheAsnAlaGlyHleAlaGlySerValAsnValAlaGlyPheSerThrIleVal 52
QY 163 AAGCGAAGTGTGCAACAGACAAAGGTGAATTAATACAGAGCTCATCCAGCTTACGGGAA 222
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 ArgAlaGAlaGAlaLysGlyAlaMetGlyLeuGlnHleAlaIleValProAsnAlaGluLeuArg 72
QY 223 CATAGAGTTGACATTGATTGACAGTCAAGAGGTGTAGTATACATCAAGCTCCCAAGAT 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GlyArgLeuLeuAlaGlyAlaTyHleAlaValValLeuLeuAspGluArgSerAlaSer 92
QY 283 GTTGCTCTCTCTCTGACAGCTGTTTCTCACTGTAATCTTGCGTAAGCTG-----GAG 336
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 93 LeuAspGlyAlaLysArgAspGlyThrLeuAlaLeuAlaAlaGlyAlaLeuCysArgGlu 112
QY 337 AAGAGCTTCACTGCTGCTGACAGTCTGCTGAGGGGTTTGCTGAGTCTCTGTTGTTTC 396
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 AlaArgSerThrGlnValPhePheLeuGlnGlyGlyGlnAlaPheSerAlaSerCys 132
QY 397 CCGGCTCTGCTGTAAGAAATCCACTAGTCCCTACCTGCAATTTCTGAGCT----- 450
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 ProGluCysSerLysGlnSerThr-----ProThrGlyLeuSerLeuProLeuSer 150
QY 451 -----TGCTTACCTGTTGCCAACATT----- 471
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151 ThrSerValProAspSerAlaGluSerGlyCysSerSerCysSerThrProLeuTyArgP 170
QY 472 -----GGGCAACCCGAATTTCTCCCAATCTTATCTTGCTGCGCAGAGAGTCTCTC 525
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 GlnGlyGlyProValGluLeuLeuSerPheLeuTyArgLeuGlySerAlaTyHleAlaSer 190
QY 526 AACAGAGCTGTATACAGAGAAATGGAGTGTATGTTTAATGCCACTATACCTGT 585
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 ArgLysArgMetLeuAspAlaLeuGlyLleThrAlaLeuIleAsnValSerAlaAsnCys 210
QY 586 CCAAGCTGACTTATCCCGAGTCTGATTTCTGCGTGGCTGGTGAATGACAGCTTT 645
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 ProAsn-----HisPheGlnGlyHleTyArgLysSerIleProValGluAspAsnHis 229
QY 646 TGTGAGAAAATTTTCCCGTGTGCAAAATCAGTATGATTTCATTGAGAAAGCAAGCC 705
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 230 LysAlaAspLleSerSerIlePheAsnGlnAlaIleAspPheIleAspSerIleTyArgP 249
QY 706 TCCAAATGATGTTTCTAGTCACTGTTTAACTGGAGATCCCGGCTCCGCAACATCGCT 765
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 AlaGlyGlyArgValPheValHleCysGlnAlaGlyLleSerAlaSerAlaThrIleCys 269
QY 766 ATCGCTCATCATCAAGAAAGATGACATGCTTTAGATAGAAGCTTACAGATTGTGA 825
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 LeuAlaTyLeuMetArgThrAsnArgValLysLeuAspGlnAlaPheGlnPheValLys 289
QY 826 GAAAAAAGACTTATATATCTCAAACTTCAATTTTGGGCAACTCTCGACATATGAG 885
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 GlnArgArgSerIleIleSerProAsnPheSerPheMetCysGlnLeuGlnPheGln 309
QY 886 AAGAAGTTTAAGAACGAGCTGAGCATCAGGCCCAAGCAAACTCAAGCTGCTGAC 945
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 310 SerGlnVal-----LeuAlaProHleCysSer----- 318

```

```

QY 946 CTGAGAAAGCCAAATGAACTGTCCCTGCTGTCTTCAGAGGTGACAAAAAGCAGACG 1005
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 319 ---AlaGluAlaGlySerProAlaMetValAlaLeuAspArgGlyThrSerThr 337
QY 1006 CCCCTCAGTCCACCCCTGCGCCGACTGTGCTACCTCAAGAGCAGAGCAAAAGCCCGT 1065
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 ValPheAsnPheProValSer-----IleProVal 347
QY 1066 CATCCCGCCACG 1077
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 HisProThrAsn 351

RESULT 14
DUS1_RAT
ID DUS1_RAT STANDARD; PRT; 367 AA.
AC 064623:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
DE CL100).
DE DUSP1 OR CL100.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Lung;
RA Muda M., Schlegel W., Arkinetal S.;
RL Submitted (JAN-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
CC kinase ERK2 on both Thr-183 and Tyr-185.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X84004; CNA58828.1; -.
DR PIR; S52265; S52265.
DR HSSP; Q16828; IMKP.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKDPHPRASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
FT HYDROLASE; Cell cycle
FT DOMAIN 20 137
FT 175 367 RHODANES.
FT ACT_SITE 258 258 PROTEIN-TYROSINE PHOSPHATASE.
FT SEQUENCE 367 AA; 39541 MW; 5112ADP290499139 CNG64;
SIMILARITY).

```


Alignment Scores:

Pred. No.: 3,33e-19 Length: 367
 Score: 423.00 Matches: 115
 Percent Similarity: 46.05% Conservative: 54
 Best Local Similarity: 31.34% Mismatches: 148
 Query Match: 11.70% Indels: 50
 Gaps: 9

US-10-029-345a-108_copy_538_2532 (1-1995) x DUS1_RAT (1-367)

```

QY 43 TTGGTGGCTCTGCTGGAAAGTGGAAAGGAAAGTCTGCTAATTGATCCGCGCATTT 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 LeuArgAlaLeuLeuArgGluArgAlaAlaGlnCysLeuLeuLeuSerCysArgSerPhe 32

QY 103 GTGGAATACATATCATCCCATTTTGGAGCCATTATATATCATCTGCTCCAGCTTATG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 PheAlaPheAsnAlaGlnHisLeuAlaGlnCysLeuLeuLeuSerCysArgSerPhe 52

QY 163 AAGCGAAGTTTGCNA-----CAGGACAAAGTTTAAATTCAGAGCTCATC 207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 ArgArgArgAlaArgGlyAlaMetGlyLeuGlnHisLeuAlaProanthrGlnLeu--- 71

QY 208 CAGCATTACGCGAAACATTAAGTTGACATTGATTCAGTCAGAGAGTTGATTGAT 267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72 -----ArgGlyArgLeuLeuAlaGlyAlaTyrHisAlaValLeuLeuAsp 87

QY 268 CAAGCTCCCAAGATGTGCTCTCTCTTCAGACTGTTTCTCATCTGATCTTGCT 327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88 GluArgSerAlaAlaLeuAspGlyAlaLeuArgSerPheLeuAlaLeuAlaGly 107

QY 328 AAAGCT-----GAGAAGACTTCATCTGTTTCAGCTGCGAGGTGGTTGCTGAG 381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 AlaLeuCyArgGlnAlaArgSerThrGlnValPhePheLeuGlnGlyGlyTyrGlnAla 127

QY 382 TTCTCTGTTGTTTCCCTGCGCTCTGTGAAGAAA----- 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 PheSerAlaSerCyPProGlnLeuCySerGlyGlnSerThrProMetGlyLeuSerLeu 147

QY 418 -----TCCACTCTAGTCCCTTACTGATTTCTCAGCTTCTTACTGTTGCCAACATT 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148 ProLeuSerThrSerValProAspSerAlaGlnSerGlyCySerSerCySerThrPro 167

QY 472 -----GGGCGCAACCGCAATTTCTTCCCAATTTTATTTGCTGCCAGCGA 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 LeuTyrAspGlnGlyGlyProValGlnLeuLeuSerPheLeuTyrLeuGlySerAlaTyr 187

QY 517 GATGTCCTCAACAAGAGCTGATACAGCAAGATGGAGTTGATTATGTTAAATGCCAGC 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 188 HisAlaSerArgIysAspMetLeuAspAlaLeuGlyIleThrAlaLeuIleAsnValSer 207

QY 577 TATACCTGTCCAAAGCTGACTTATCCCGAGTCTCATTTCCCTGCGTGCCTTGGAAT 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 AlaAsnIysPheAsn---HisPheGlnGlyHisTyrGlnTyrIysSerAlaIleProValGln 226

QY 637 GACAGCTTTTGTGAGAAATTTTCCGGTGGTGGCAAAATCAGATGATTCATTCAGAAA 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 AspAsnHisIysAlaAspIleSerSerThrPheAsnGlnAlaIleAspPheIleAspSer 246

QY 697 GCAAAAGCTCCAAATGAGTGTTCATAGTCACATGTTTACCTGGAGATCTCCGCTCGCC 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 IleIysAspAlaGlyArgValPheValHisCysGlnAlaGlyIleSerArgSerAla 266

QY 757 ACCATGCGTATCGCTTCATCATGAGAGAGATGACATGCTTTCATGTAAGCTTACAGA 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 ThrIleCysLeuAlaTyrIleuMetArgThrAsnArgValIysLeuAspGlnAlaPheGln 286

QY 817 TTGTGAAAAGAAAAAAGCTATATATTCAAAATTCATTTTGGGCGCAATCTCG 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 287 PheValIysGlnAlaArgSerIleIleSerProAsnHisPheMetGlyGlnLeu 306

QY 877 GACTATGAGAAGAGATTAAAGAACAGACTGAGCATCAGGCGCAAGCAAACTCAAG 936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 GlnPheGlnSerGlnVal-----LeuAlaProHisCysSer----- 318
  
```

```

QY 937 CTGCTGACCTGGAGAGCAAAATGAACTGTGCTGCTCTCTCAGAGGGTGCAGAGAA 996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 319 -----AlaGlnAlaGlySerProAlaMetAlaValLeuAspArgGlyThrSer 334

QY 997 AGCGAGAGCGCCCTCAGTCCACCTGTGCGCACTGTCTTACTCAGAGCGAGAGACAA 1056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 ThrThrThrValPheAsnPheProValSer----- 344

QY 1057 AGGCGCGTGCATCCCGCCAGC 1077
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 345 IleProValHisPheThrAsn 351

RESULT 15
DUS1_HUMAN
ID DUS1_HUMAN STANDARD; PRT; 367 AA.
AC P28562;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 1 (PC 3.1.3.48) (PC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
DE CL100) (Dual specificity protein phosphatase hvh1).
GN DUSP1 OR PTPN10 OR MKP1 OR CL100 OR VH1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Forebrain;
RX MEDLINE=93024952; PubMed=1406996;
RA Keyes S.M., Emalie E.A.,
RT "Oxidative stress and heat shock induce a human gene encoding a
RT protein-tyrosine phosphatase."
RL Nature 359:644-647 (1992).
RN 12;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo W.F., Caesavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallat S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grilwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schermer A., Schein J.E., Jones S.J.M., Maira W.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
CC kinase ERK on both Thr-183 and Tyr-185.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- INDUCTION: By oxidative stress and heat shock.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
  
```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
CC or send an email to license@ebi.ac.uk).

DR EMBL; X68277; CAA48338.1; -.
DR EMBL; BC022463; AAH22463.1; -.
DR PIR; S29090; S29090.
DR HSP; Q16828; IMKP.
DR Gene; H0003064; DUSP1.
DR MIM; 600714; -.
DR GO; GO:0004726; F:non-membrane spanning protein tyrosine phos. .; TAS.
DR GO; GO:0006979; P:response to oxidative stress; TAS.
DR InterPro; IPR000340; Ds phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DUSPc; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DUSPc; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANESE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase; Cell cycle.
FT DOMAIN 20 137 RHODANESE.
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE (BY
SIMILARITY).
SQ SEQUENCE 367 AA; 39297 MW; 11BD1D39A9FCD51F CRC64;

Alignment Scores:
Pred. No.: 5.14e-19 Length: 367
Score: 420.00 Matches: 113
Percent Similarity: 46.69% Conservative: 56
Best Local Similarity: 31.22% Mismatches: 159
Query Match: 11.62% Indels: 34
DB: 1 Gaps: 7

US-10-029-345A-108_COPY_538_2532 (1-1995) x DUS1_HUMAN (1-367)

QY 43 TTGGTGGCTCTGCTGGAAGTGAACGGAAGGCTGTAATTGATAGCGGCAATT 102
DB 13 LeuArgAlaLeuLeuGlyGluArgAlaIaIaGlnCysLeuLeuLeuAspCysArgSerPhe 32
QY 103 GTGGAATACATATACATCCCACTTTTGAAGCCATTATATCACTGCTCCAGCTTATG 162
DB 33 PheAlaPheAsnAlaGlyHisIleIaGlySerValaAsnValaArgPheSerThrIleVal 52
QY 163 AAGCGAAGTTCACACAGACGAAAGTAAATTATACAGAGCTCATCCAGATTCACGCAAA 222
DB 53 ArgArgArgAlaIaGlyAlaIaMetCysLeuGluHisIleValProAsnAlaGluLeuArg 72
QY 223 CATAGAGTTGACATTGATTGACAGCAAGAGCTTTAGTTTACGATCAAGCTCCCAAGAT 282
DB 73 GlyArgLeuLeuAlaGlyAlaIaTyriIaIaValIaLeuLeuAspGluArgSerAlaIa 92
QY 283 GTTGCTCTCTCTCTTCAAGACTGTTTCTCACTGACTTCTGGGTAAATG-----GAG 336
DB 93 LeuArgGlyAlaIaIaArgArgPglYThrLeuAlaLeuAlaIaGlyAlaLeuCysArgGlu 112
QY 337 AAGAGCTTCACTGCTGCTGAGAGTGTGAGGTGGTTTGAGTGTCTCTGTTGTTTC 396
DB 113 AlaArgAlaIaIaGlnValaPhePheLeuLysGlyGlyTyriGluAlaPheSerAlaSerCys 132
QY 397 CCTGGCTCTCTGTAAGAAA-----TCACTCTTA 426
DB 133 ProGluLeuCysSerIaIaSerIaSerThrProMetGlyLeuSerLeuProLeuSerThrSer 152
QY 427 GTCCCTACCTGCAATTTCTACGCTTGCTTACTGTTGCCACATT----- 471

DB 153 ValProAspSerAlaGlySerGlyCysSerSerCysSerThrProLeuTyriAspGlnGly 172
QY 472 GGGCGCAACCGGAATTTCTTCCCAATCTTATCTTGAGCTCCAGCGAGATGTCTCAACAG 531
DB 173 GlyProValGluIleLeuProPheLeuTyriLeuGlySerAlaTyriHisAlaSerArgLys 192
QY 532 GAGCTGATACAGCAAGAAATGGATTTGTTATGTGTTAAATCCAGCTTACTCTCCAAAG 591
DB 193 AspMetLeuAspAlaLeuGlyIleThrAlaLeuIleAsnValSerAlaAsnCysProAsn 212
QY 592 CCTGACTTATCCCGAGTCAATTCCTCGTGTGCTGATGATGACAGCTTGTGAG 651
DB 213 ---HisPheGluGlyHisTyriGlyTyriLysSerIleProValGluAsnHisLysAla 231
QY 652 AAAATTTGGCGGTGGTGGCAAAATCAATGATTTCAATTGAGAAAGCAAAAGCTCCAAAT 711
DB 232 AspIleSerSerThrPheAsnGluAlaIleAspPheIleAspSerIleLysAsnAlaGly 251
QY 712 GAGATGTTCTAGTGCACTGTAGCTGGGATCTCCGCTCCGCAACATCGCTATGCGC 771
DB 252 GlyArgValPheValHisCysGlnAlaGlyIleSerArgSerAlaThrIleCysLeuAla 271
QY 772 TACATCATGAAGAGATGACATGTCTTATGATGAGCTTACAGATTTGTGAAGAAA 831
DB 272 TyriLeuMetArgThrAsnArgValLysLeuAspGluAlaPheGluPheValLysGlnArg 291
QY 832 AGACCTTACTATCTCCAACTTCAATTTCTGGCGCAACTCTGATGAGAGAG 891
DB 292 ArgSerIleIleSerProAsnPheSerPheMetCysGlnLeuLeuGlnPheGluSerGln 311
QY 892 ATTAAGACCAAGACTGAGATCAGGAGCCCAAAAGCAAACTCAAGCTGCTGACCTGGAG 951
DB 312 Val-----LeuAlaProHisCysSer-----Ala 319
QY 952 AAGCGAATGAACCTGTCCCTGCTGTCTCAGAGGAGTGAACGAGAGAG----- 1005
DB 320 GluAlaGlySerProAlaMetAlaValLeuAspArgGlyThrSerThrThrValPhe 339
QY 1006 -----CCCTCAGTCCAGCCCTGTGCGCACTGTGCTTACCTCAGAGCAAGCAAGAG 1059
DB 340 AsnPheProValSerIleProValHisSerThrAsnSerAlaLeuSerTyriLeuGlnSer 359
QY 1060 CCGGTG 1065
DB 360 Profile 361

Search completed: June 21, 2004, 12:34:52
Job time : 42.9747 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: June 21, 2004, 12:28:25 ; Search time 103.154 Seconds
(without alignments)
12204.244 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 3615
Sequence: 1 atgagccatgagatgattg9.....tggaaatcatctgagctctcc 1995

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+npz.model -DEV=xlp
-Q/cg21/USPTO.spool.p/US10029345/runat.21062004.122816.4193/app.query.fasta.1.3278
-DB=SPREMBL.25 -QFMT=fastaan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45
-DOCNALIGN=200 -THR_SCORE=pcpr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcpr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10029345.QCGN.1.1.168.0/runat.21062004.122816.4193 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NRG_SCORES=0 -WAIT -DSPELLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3399	94.0	665	4 Q96N49	Q96N49 homo sapien

2	3379.5	93.5	662	4 Q81VT8	Q81VT8 homo sapien
3	3163	87.5	616	4 Q8N5T1	Q8N5T1 homo sapien
4	3079.5	85.2	660	11 Q920R2	Q920R2 mus musculus
5	2954.5	81.7	677	11 Q99MG6	Q99MG6 mus musculus
6	2613	72.3	622	11 Q99MG5	Q99MG5 mus musculus
7	1455.5	40.3	355	11 Q9AG16	Q9AG16 mus musculus
8	1327	36.7	625	4 Q86SS8	Q86SS8 homo sapien
9	1304	36.1	665	11 Q7RSZ9	Q7RSZ9 mus musculus
10	1114	30.8	300	11 Q9AG15	Q9AG15 mus musculus
11	891	24.6	206	11 Q8B2E4	Q8B2E4 mus musculus
12	653	18.1	143	4 Q96Q52	Q96Q52 homo sapien
13	628.5	17.4	657	5 Q8ST19	Q8ST19 caenorhabdi
14	515	14.2	606	5 Q8ST18	Q8ST18 caenorhabdi
15	487.5	13.5	367	13 Q7S2F3	Q7S2F3 brachydanio
16	476	13.2	439	5 Q81G35	Q81G35 caenorhabdi
17	469	13.0	382	13 Q7T2L9	Q7T2L9 gallus gall
18	469	13.0	483	11 Q8R3L3	Q8R3L3 mus musculus
19	449	12.4	383	13 Q8UM48	Q8UM48 fugu rubrip
20	447	12.4	411	4 Q13649	Q13649 homo sapien
21	445	12.3	368	4 Q8NFX0	Q8NFX0 homo sapien
22	439.5	12.2	388	11 Q8BFX3	Q8BFX3 mus musculus
23	437	12.1	382	13 Q7T2L8	Q7T2L8 brachydanio
24	433.5	12.0	369	13 Q91790	Q91790 xenopus lae
25	431.5	11.9	369	13 Q90W58	Q90W58 xenopus lae
26	431	11.9	362	13 Q803B2	Q803B2 brachydanio
27	427.5	11.8	378	13 Q91663	Q91663 xenopus lae
28	424	11.7	367	11 Q64193	Q64193 rattus sp.
29	421	11.6	367	11 Q63683	Q63683 rattus norv
30	415.5	11.5	353	13 Q42253	Q42253 gallus gall
31	415	11.5	360	13 Q7ZVL8	Q7ZVL8 brachydanio
32	413	11.4	318	11 Q80ZM1	Q80ZM1 mus musculus
33	412	11.4	436	11 Q99KC2	Q99KC2 mus musculus
34	406	11.2	452	11 Q8KLS9	Q8KLS9 mus musculus
35	406	11.2	452	11 Q7TNL7	Q7TNL7 mus musculus
36	391.5	10.8	354	4 Q8N4A4	Q8N4A4 homo sapien
37	353	9.8	411	5 Q95SV1	Q95SV1 drosophila
38	353	9.8	411	5 Q9VW55	Q9VW55 drosophila
39	345	9.5	303	4 Q9NSM1	Q9NSM1 homo sapien
40	333	9.2	476	5 Q46122	Q46122 drosophila
41	328	9.1	476	5 Q9VW8	Q9VW8 drosophila
42	306	8.5	177	11 Q9CSL5	Q9CSL5 mus musculus
43	292.5	8.1	359	5 Q44128	Q44128 caenorhabdi
44	287	7.9	1045	5 Q9NKX1	Q9NKX1 drosophila
45	287	7.9	1193	5 Q8IMW8	Q8IMW8 drosophila

ALIGNMENTS

RESULT 1
ID Q96N49 PRELIMINARY; PRT; 665 AA.
AC Q96N49;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ131411.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsui T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Nagasawa M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Negahari K., Maeno Y., Nagai K.,
RA Isegai T.,
RA "MEDO human cDNA sequencing project";
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK05973; BAB71060.1; -

QY 1741 GCCTAGAGCTGACGACGCTGCCACTTGCGAGACCAAGCTTATTCTGCGCAGGCGG 1800
DB 581 AATySerCySerGlnLeuProthrCySGlyAspGlnAlaTySerValAlaArgArg 600
QY 1801 CAGAACCAAGTACAGAGCTGACTCGCGGAGAGCTGCGATGAGAGAGCCCTTTGAA 1860
DB 601 GlnLysProSerAspArgAlaAspSerArgSerTrpHisGlnGlnSerProPheGln 620
QY 1861 AAGCAGTTTAAACGACAGAGCTGCCAATGGAATTGGAGAGACATCATGTCAAGAAC 1920
DB 621 LysGlnPheLysArgArgSerCySGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGCTCAGCGGAAGAGCTGGGGAAGAGCGAGTCAGCTTCTTCCGCGACATGGA 1980
DB 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCATGAGGCTCTCC 1995
DB 661 IleIleGlnValSer 665

RESULT 2

Q81VT8 PRELIMINARY; PRT; 662 AA.

AC 081VT8;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Similar to dual specificity phosphatase 16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Duoenum;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC042101; AAH42101.1; -
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:006470; P:Protein amino acid dephosphorylation; IEA.
DR InterPro; IPR008340; DS phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00782; DSpC; 1.
DR PRINTS; PR01764; MAPKDPHPTAS.
DR SMART; SM00195; DSpC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANASE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 662 AA; 72818 MW; 350534ER0652B98F CRC64;

Alignment Scores:

Pred. No.: 7.09e-251 Length: 662
Score: 3379.50 Matches: 660
Percent Similarity: 99.40% Conservative: 1
Best Local Similarity: 99.25% Mismatches: 3
Query Match: 93.49% Indels: 3
DB: 4 Gaps: 1

US-10-029-345a-108_copy_538_2532 (1-1995) x Q81VT8 (1-662)

QY 1 ATGGCCATGAGATGATTGGAACCAATTGTTACTGAGAGGTTGCGCTGCTGCGAA 60
DB 1 MetAlaHisGlnMetCilleglyThrGlnIleValAlaThrGlnArgLeuValAlaLeuGln 20
QY 61 AGTGAACGAGAAAAGTGTCTTAATTGATGACGCGCATTTGTGAATACAATACATCC 120
DB 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnLysPheThrSer 40

QY 121 CACATTTGGAAGCCATTATATCACTGTCCCAAGCTTATGAGAGGAAGTTGCAACAG 180
DB 41 HisIleLeuGlnAlaIleAsnIleAsnCySerIleLeuMetLysArgArgLeuGlnGln 60
QY 181 GACAAAGTTTATTACAGAGCTCATCAGACATTCAGCAAAACATAGTTGACATTGAT 240
DB 61 AspLysValLeuIleThrGlnLeuIleGlnIleSerAlaLysHisLysValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGTGATTTCATCAAGTCCCAAGATTTGCTCTCTCTCA 300
DB 81 CySerGlnLysValValAlaLysArgGlnSerSerGlnAspValAlaSerLeuSer 100
QY 301 GACTGTTTCTACCTGACTCTGGGTAACTGAGAGAGCTTCACTGTCTCACTG 360
DB 101 AspCySPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGTGGTTTGTGCTGAGTTCTCGTGTGTTTCTCGGCTCTGTGAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCySPheProGlyLeuCySGlnGlyLysSer 140
QY 421 ACTGATGCTTACCTGACATTTCTGAGCTTGTCTTACCTGTTCCCAACATGGCCAAC 480
DB 141 ThrLeuValProThrCySPheSerGlnProCySPheProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTATCTTGCTGCGCAGCGAGATGCTCTCAACAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuLysLeuGlyCySGlnArgAspAlaLeuAsnLysGlnLeuMet 180
QY 541 CAGCAGAAATGGAGTTGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyLysGlyLysValLeuAsnAlaSerAsnThrCySPheProAspPhe 200
QY 601 ATCCCGAGCTCATTTCTGCGCTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCySGlnLysIleLeu 220
QY 661 CGGTGTTGGACAAATCAGTAGATTTCATTAGAGAAAGCAAAAGCTCCAAATGAGTGTG 720
DB 221 ProThrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCyVal 240
QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCCACATCGCTATCGCTATCATGATG 780
DB 241 LeuValHisCyLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaLysIleMet 260
QY 781 AAGAGATGAGACATGCTCTTAAAGTGAAGCTTAAAGATTGTGAAGAAAAAGACTACT 840
DB 261 LysArgMetAspMetSerLeuAspGlnAlaLysArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGGAATGAGAGAAAGATTAAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysIleLysAsn 300
QY 901 CAGACTGAGACATCAGGGCCAAAGACCAACTCAAGCTGTGACCTGTGAAGACCAAT 960
DB 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLeuLeuHisIleGlnLysProAsn 320
QY 961 GAACCTGTCTCTGTCTGACAGGCTGACAGAGAAAGCAGACGCCCTCATCTCAAGCC 1020
DB 321 GlnProValProAlaValSerGlnGlyGlnLysSerGlnLysProLeuSerProPro 340
QY 1021 TGTGCGGACCTGCTACCTCAGAGGAGCAGCAAGAGCCGCTGATCCGCGACGGTG 1080
DB 341 CyAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAla----- 358
QY 1081 CCCAGCGTCCACGCTGACGCGCTGCTGTTAGAGACAGCCGCTGTATCAGCGCTC 1140
DB 359 ---SerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 377
QY 1141 AGTGGCTGACCTGTGCGCAGACAGGCTGGAAAGACAGCAATAGCTCAAGCTTCTTTC 1200
DB 378 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLysArgSerPhe 397
QY 1201 TCTCTGATATCAAAATCAGTTTATGATGACGAGATGGCAGCATCTTCAATGAGCTTC 1260

```

Db      398 SerLeuAspIleIysSerValSerIlySerIalSerIalSerLeuHISgIlyPhe 417
Qy      1261 TCCCATATGAGAAAGTCTTTGGAATACTTACAAACCTTCCACTACTCTGGATGGGACAC 1320
Db      418 SerSerSerGluAspAlaLeuGluTyrIlyProSerThrThrLeuAspGlyThrIbn 437
Qy      1321 AAGCTATGCGAGTCTCCCTGTTGACAGACTTTCGAGAGACATCCCGAAACAGATCT 1380
Db      438 LysLeuCySgInPheSerProValGlnGluLeuSerGlnGlnThrProGluThrSerPro 457
Qy      1381 GATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db      458 AspYsgIuGluAlaSerIleProIlyLysLeuGlnThrAlaArgProSerIbn 477
Qy      1441 AGCAAGCGATTGATTCGATCGATCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db      478 SerIlyArgLeuHISerValAlaGlnThrSerSerIlyThrAlaGlnArgSerLeuLeu 497
Qy      1501 TCTCCATGATGATGAAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db      498 SerProLeuHISerIlySerValGluAspAenIlyrHISerIbnPheLeuPheGly 517
Qy      1561 CTTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
Db      518 LeuSerThrSerGlnGlnIleuThrIlySerIalSerIalSerIalSerIalSerIalSer 537
Qy      1621 TGGATATCTTGGCCCCCAAGCTCTACCCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db      538 SerAspIleLeuAlaProGlnThrSerThrProSerLeuHISerIlyrPheAla 557
Qy      1681 ACAGAGCTCTGACACTTCTACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db      558 ThrGlnSerSerHISerIlySerIalSerIalSerIalSerIalSerIalSerIalSer 577
Qy      1741 GCTTACAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db      578 AlaIlySerCysSerGlnLeuProIlyrHISerIalSerIalSerIalSerIalSerIalSer 597
Qy      1801 CAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db      598 GlnIlyProSerAspAlaSerIalSerIalSerIalSerIalSerIalSerIalSerIalSer 617
Qy      1861 AAGCAGTTTAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db      618 LysGlnPheIlyrArgIlySerCysGlnMetGluPheGlyGlnSerIleMetSerGluAsn 637
Qy      1921 AGGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
Db      638 ArgSerArgGlnGluLeuGlyValGlySerGlnSerSerPheSerGlySerMetGlu 657
Qy      1981 ATGATTGAGGTCTCC 1995
Db      658 ILeIleGluValSer 662

```

```

DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro: IPR000340; DS phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR000387; Tyr_phosphatase.
DR Pfam: Pf00782; DSPC; 1.
DR PRINTS: PR01764; MAPKPHPTASE.
DR SMART: SM00195; DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS00054; TYR_PHOSPHATASE_DUAL; 1.
DR Hypothetical protein.
FT NON TER
SQ
SEQUENCE 616 AA; 67636 MW; 2CBOB14482F2AD72 CRC64;

Alignment Scores:
Pred. No.: 2,96e-234 Length: 616
Score: 3163.00 Matches: 613
Percent Similarity: 99.84% Conservative: 2
Best Local Similarity: 99.51% Mismatches: 1
Query Match: 87.50% Indels: 0
DB: 4 Gaps: 0

US-10-029-345a-108_copy_538_2532 (1-1995) x Q8N5T1 (1-616)
Qy      148 TGCCTCAAGCTTATGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 207
Db      1 CysSerIlySerLeuMetIlyrArgIlyLeuGlnGlnAspIlyValIleuIleThrGluLeuIle 20
Qy      208 CAGCATTCAGCGAAACATTAAGGTTGACATTTGACATTTGACATTTGACATTTGACATTTGACAT 267
Db      21 GlnHISerAlaIlyHISerIlyValAlaAspCysSerGlnIlyValIleuIlyrAsp 40
Qy      268 CAAAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 327
Db      41 GlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrValIleuLeuGly 60
Qy      328 AAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 387
Db      61 LysLeuGlnIlySerPheAsnSerValHISerLeuLeuAlaIlyGlyPheAlaGluPheSer 80
Qy      388 CGTTGTTTCCCTGGGCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 447
Db      81 ArgCysPheProGlyLeuCySgIuGlyLysSerThrLeuValProThrCysIleSerGln 100
Qy      448 CTTGCTTACTGTTGCCAATTTGGGCCAACCCGAATCTTCCCAATCTTATCTTGGC 507
Db      101 ProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuIlyrLeuGly 120
Qy      508 TGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 567
Db      121 CysGlnArgAspValAlaLeuAsnIlyGlnLeuMetGlnGlnAsnGlyIleGlyTyrValLeu 140
Qy      568 AATCCAGCTTAACTGTTCCAAAGCCTGACTTATCCCGAGTCCATTTCTGCTGCTG 627
Db      141 AsnAlaSerAsnThrCysProIlyProAspPheIleProGlnSerHISerPheLeuAlaVal 160
Qy      628 CTTGTAATGACAGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 687
Db      161 ProValAsnAspSerPheCysGlnIlySerIleLeuProThrLeuAspIlySerValAspPhe 180
Qy      688 ATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 747
Db      181 ILeGluIlyrAlaIlyAlaSerAsnGlyCysValLeuValHISerLeuAlaGlyIleSer 200
Qy      748 CGCTCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 807
Db      201 ArgSerAlaThrIleAlaIleAlaIlyrIleMetIlyrArgMetAspMetSerIleuAspGlu 220
Qy      808 GCTTACAGATTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 867
Db      221 AlaIlyrArgPheValIlyGlnIlyrArgProThrIleSerProAsnPheAsnPheLeuGly 240

```


QY 868 CAATCTGAGCTAGTGAAGAGATTAGAACGAGCTGAGAGATCAGGGCCAAAGAGC 927
DB 241 GlnleuenuaprytGluLysLysIleYenahmglnthrgYalaserGlyProLysSer 260
QY 928 AAATCAAGCTGTGACACTGTGAGAGAGCAAAATGACCTGTCTCCGTCTCAGAGGCT 987
DB 261 LysleuenuYleuenuHsleuGluYerProahndiuroValProAlaValaserGluY 280
QY 988 GAGACAGAAAAGCGAGAGCGCCCTCAGTTCACCTGTGCGACCTGTCTGACTCAGAGGCA 1047
DB 281 GlyGlnYserSerGluThrProLysSerProProCysAlaApsSerAlaThrSerGlnAla 300
QY 1048 GAGGAGCAAAAGCGCGGATCCCGCAGAGGTGCGAGCGGCGCCAGCGAGCGCGCTG 1107
DB 301 AlaeGlylnmrgProValHsProAlaserValProserValProserValGlnProSer 320
QY 1108 CTGTGAGAGACAGCCCGCTGTGTACAGCGCTCAGTGGGCTGACCTGTGCGAGAGAG 1167
DB 321 LeuenuGluApsSerProleuValGlnAlaLeuSerGlyleuHsleuSerAlaApsArg 340
QY 1168 CTGGAGACAGCAATTAAGCTCAGAGCGTTCTCTCTCGATATCAATCAGTTTCATAT 1227
DB 341 LeuGluApsSerAenLysleuLysArgSerPheSerleuApsIleLysSerValSerTyr 360
QY 1228 TCAGCGCAGATGGAGATCCTTACATGGCTTCCTCATCAGAGAGATGGCTTGGAAATAC 1287
DB 361 SerAlaserMetAlaAlaserLeuHsGlyPheSerSerSerGlnApsAlaLeuGlnTyr 380
QY 1288 TACAAACCTTCACTACTGTGATGAGAGCAACAAGATATGCCAGTTCTCCCTGTTCAG 1347
DB 381 TyrlYserProSerThrThrleuApsGlyThrAenLysleuYsglnPheSerProValGln 400
QY 1348 GAATATCGAGAGACTCCCGAAACAGTCTGTATAGAGAGAGCGAGCATCCCAAG 1407
DB 401 GluleuSerGluGlnThrProGluThrSerProApsYsglnGlnAlaserIleProLys 420
QY 1408 AAGGTGAGACCGCGAGCGCTTCAGACAGCGAGAGGATTTGGTGGTCAGAAC 1467
DB 421 LysleuGlnThrAlaArgProSerApsSerleuYsArgleuHsleuSerValArgThr 440
QY 1468 AGCAGCTGAGACCGCGCAGAGAGTCCCTTTATCTCAGTGCATCGAAGTGGAGCGTG 1527
DB 441 SerApsSerGlyThrAlaGlnArgSerleuLeuSerProleuHsAgsSerGlySerVal 460
QY 1528 GAGGACATTAACACACAGCTTCCTTTGGCGCTTTCACACAGCCAGACAGCTCAGC 1587
DB 461 GluApsAenLysIleThrSerPheleuPheGlyleuSerThrSerGlnGlnHsleuThr 480
QY 1588 AAGCTGCTGGCGCGCTTAAAGGCTGCACTCGGATATCTTGGCGCCCGCAGACTCT 1647
DB 481 LysSerAlaGlyleuGlyleuLysGlyIlePheHsSerApsIleleuAlaProGlnThrSer 500
QY 1648 ACCGCTTCCGTGACAGCAGCTGATATTTGGCAGAGAGTCCCTCAGACTTCTACTGCGC 1707
DB 501 ThrProSerleuThrSerSerTyrPheAlaThrGluSerSerHsIlePheIleYserAla 520
QY 1708 TCAGCCATCTTACGAGGAGCTGCGATTACTGTGCTTACAGCTGACCGACAGCTGCCACT 1767
DB 521 SerAlaIleYrGlySerAlaserTyrSerAlaIleYserCysSerGlnleuProThr 540
QY 1768 TGGCGAGACCAAGCTATTTCTGTGGCGAGCGGAGAGAACCAATGACAGAGCTGACTCG 1827
DB 541 CysGlyApsGlnValIleYserValArgArgGlnYsProSerApsArgAlaApsSer 560
QY 1828 CGGCGAGCTGAGTGAAGAGAGCCCTTTGAAAAGAGATTAAACGAGAGCTGCCAA 1887
DB 561 ArgArgSerTrpHsGluGlnSerProPheGluYsglnPheYsAgsArgSerCysGln 580
QY 1888 ATGGAATTTGAGAGAGCATCATGTCAAGAAACAGGTTCAGCGGAAAGCTGGGGAAAGTG 1947
DB 581 MetGluPheGlyGlnSerIleMetSerGluAenAgsSerArgGlnGluLeuGlyVal 600
QY 1948 GCGAGTACGTCTAGCTTTTCGGGAGAGATGAATCATTTAGAGTCTCC 1995

DB 601 GlySerGlnSerPheSerSerGlySerMetCgluIleleuValSer 616
RESULT 4
ID Q920R2 PRELIMINARY; PRT; 660 AA.
AC Q920R2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE MAP kinase phosphatase-7.
GN DUSP16 OR 3830417M17RIK OR MKP-7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
RT "MKP-7, a Novel Mitogen-activated Protein Kinase Phosphatase,
Function as a Shuttling Protein."
RL J. Biol. Chem. 276:39002-39011(2001).
DR EMBL: AB052157; BAB47240.1; -.
DR HSSP: Q16828; 1MKP.
DR MCD; MG1:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phospn.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MARKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PSS0206; RHODANSE_3; 1.
DR PROSITE; PSS0383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PSS0056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SQ SEQUENCE 660 AA; 72695 MW; DB609FCDADA4A309 CRC64;
Alignment Scores:
Pred. No.: 7 82e-228 Length: 660
Score: 3079.50 Matches: 599
Percent Similarity: 94.29% Conservative: 28
Best Local Similarity: 90.08% Mismatches: 33
Query Match: 85.19% Indels: 5
DB: 11 Gaps: 3
US-10-029-345a-108_COPY_538_2532 (1-1995) x Q920R2 (1-660)
QY 1 ATGGCCCATGAGATGATGGAGTCAATTTGTTACTGAGAGGTTGGCTGTGGGAA 60
DB 1 MetAlaHsleuMetIleGlyThrGlnIleValThrGluSerleuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAAAGTGCCTGATTAATGATAGCCGCGCATTTGTGGAAATCAATACATCC 120
DB 21 SerGlyThrGluYsValleuLeuIleApsSerArgProheValGlnTyrAenThrSer 40
QY 121 CACATTTGGAAGCCATTATATCACTGCTCAAGCTTATGAAGCGAAAGTTGCAACAG 180
DB 41 HsleuGluAlaIleAenIleAenCysSerIleleuMetLysArgGlnGln 60
QY 181 GACAAAGTTTAATTAACAGAGCTATTCAGCATTCAGCGAACAATAGTTGACATTGAT 240
DB 61 AspLysValleuIleThrGluLeuIleGlnHsSerAlaYsHsYsValApsIleAps 80

QY	1241	TGAGGTGAGAAAGTTGTGATTACGATCAAAAGCTCCAAAGATGTGCTCTCTCTTCA	300
Db	81	:::	300
QY	301	GACTGTTTCTCATCTGTAACTCTTGAGGTAAACGTGAAGAAAGACTTCAACTCTGTTCACCTG	100
Db	101	AspCySphLeuThrValLeuLeuLeuGlyLysLeuGlnAAsSerPheAAsnSerValHisLeu	100
QY	361	CTTGAAGAGGGGTTTGTCTGAGTTCTCTCGTTGTTCCTCGAGCTCTGTGAAGAAATGC	420
Db	121	LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuCysGlnGlyLysSer	420
QY	421	ACTTATGCTCCATCCAGCTATTTCTCAAGCTTGCTTACTGTGGCAACATTTGGCCAACT	140
Db	141	ThrLeuValProThrCysAlaSerGlnProCysLeuProValAlaAsnIleGlyProThr	140
QY	481	CGAATTTCTCCAACTTTATCTTGTGGCTGCCAGAGATGTCCTCAACAAGAGAGATGA	540
Db	161	ArgIleLeuProAsnLeuArgLysIleGlyCysGlnArgSphValLeuAsnLysAspLeuMet	540
QY	541	CAGCAGAAATGGAGATTGTATTGTGTAAAGCCAGCTATCTGTGCCAAAGCTCAACTTT	180
Db	181	GlnGlnAAsnGlyIleGlyIlyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe	180
QY	601	ATCCCCGAGTCTCAATTTCTGCTGTGCTCCGTAAATGACAGCTTTTGTGAAGAAATTTG	200
Db	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu	200
QY	661	CCGTGTGTGACAAATAGATGATTTTCATTTGAGAAAGCAAAAGCTCCAAATGATGTGT	220
Db	221	ProThrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal	220
QY	721	CTAGTACACTGTTTACGTGGATCTCCGCTCCGACATACGCTATGAGCTTAATGATAG	240
Db	241	LeuIleHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIlyrIleMet	240
QY	781	AAGAGATGACACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGACTTACT	260
Db	261	LysArgMetAspMetSerLeuAspGlnAlaIlyrArgPheValLysGlnLysArgProThr	260
QY	841	ATATCTCCAACTTCAATTTTGTGGCCAACTCTCGATGACATGAGAAAGATTTAAGAC	280
Db	281	IleSerProAsnPheAsnPheMetGlyGlnLeuMetAspIlyrGlnLysThrIleAsnAsn	280
QY	901	CAGACTGAGCATTCAGAGGCAAAAGGCAAAATCAAGCTGCTGACCTGGAAGACCAAT	300
Db	301	GlnThrGlyMetSerGlyProLysSerLysLeuLysIleLeuHisPheLysProSer	300
QY	961	GAACTGTCCTCTGCTGTCAAGAGGTGACAGAAAGACGACGCTCAGTCCACCC	1020
Db	321	GlnProValProAlaAlaSerGlnGlyIlyrLysSerAlaLeuSerLeuSerProPro	1020
QY	1021	TGTGCGCACTGTGTCACCTCAAGAGGAGAGCAAAAGCCGTGATCCGCCAGCGTG	340
Db	341	CysAlaAsnSer---ThrSerGlnAlaSerGlyGlnArgLeuValHisProAla-----	340
QY	1081	CCGAGCGTGGCCAGCGTCAAGCCGTGCTGTTAGAGACAGGCCCTGTGACAGCGCTC	357
Db	358	---SerValProSerLeuGlnProSerIleLeuGlnAspSerProLeuValGlnAlaLeu	357
QY	1141	AGTGGAGTGCACCTGTCCGACAGCAGCGTGAAGACAGCATTAAGCTCAAGCTTCTTC	376
Db	377	SerGlyLeuGlnLeuSerSerGlyLysLeuGlnAspSerThrLysLeuLysArgSerPhe	376
QY	1201	TTCCTGATATCAATCAAGTTTCATATTCAGCAGCAGACAGATCCTTAACATGCTTC	1260
Db	397	SerLeuAspIleLysSerValSerIlyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe	1260
QY	1261	TCCTCATCAGAAATGCTTTGGAATCAACAACTTCACTACTGTGATGGACCAAC	416
Db	417	---SerSerGlnGlnAlaLeuAspIlyrCysLysProSerAlaThrLeuAspGlyThrAsn	416
QY	1321	AAGCATATGCAGATTTCTCCCTGTTCAGAGAACTATCGAGAGAGACTCCGCAAAACAGATCTT	435
Db	1321	1321	1380

[illegible]

DR GO: GO:0005515; F:protein binding; IPI.
DR GO: GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPc; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS: PR01764; MAPKPHPTASE.
DR SMART: SM00195; DSpc; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANES_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrobase: Kinase.
SQ SEQUENCE 677 AA; 74550 MW; 8B6D5B7096CE2FC CRC64;

Alignment Scores:

Prod. No.:	3_15e-218	Length:	677
Score:	2954.50	Matches:	579
Percent Similarity:	92.19%	Conservative:	35
Best Local Similarity:	86.94%	Mismatches:	45
Query Match:	81.73%	Indels:	8
DB:	11	Gaps:	3

US-10-029-345A-108_COPY_538_2532 (1-1995) x Q99MG6 (1-677)

QY 1 ATGGCCATGATGATGATGGAATCAATGTTACTGAGAGGTTGCTGCTGGAA 60
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluSerLeuValAlaLeuLeuGlu 20
QY 61 AGTGAACGGAAGAAAGCTGCTTATGATGACCGCCATTTGGTAATCATCATCTCC 120
Db 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyraThrSer 40
QY 121 CACATTTGGAGCATTAATATCACTGCTCCAAAGTTTGAAGGAGGTTGCAACAG 180
Db 41 HisIleuGluAlaIleAsnIleAsnGlySerLeuMetCysArgGluGln 60
QY 181 GACAAAGTTTAATTACAGAGCTCATCAGCATTCAGCGAACAATAGTTGACATTGAT 240
Db 61 AspArgValLeuIleThrGluLeuIleGlnHisSerIleValHisIleValAsp 80
QY 241 TGCAGTCAAGAGTTGATGATTACGATCAAAAGCTCCAAAGATGTTGCTCTCTTCA 300
Db 81 CysAsnGlnIlyValIleValIlyAspGlnSerSerIleAspValGlySerLeuSer 100
QY 301 GACGTTTCTCACTGACTTCTGGGTAACCTGAGAGAAGCTTCACATCTGTTACCTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnArgSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGAGTGGGTTTGGGATTTCTCGTTGTTCCCTGGCCCTGCTGAGAGAAATTC 420
Db 121 LeuAlaGlyIleGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlySer 140
QY 421 ACTCTAGTCCCTACCTCATTTCTCAGCCCTTGCTTACCTGTTGCAACATTTGGCAACC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 GCAATTTCTCCCATCTTTACTTCTTGCTGCGCAGGAGATGCTCCCAACAGAGCTGTA 540
Db 161 ArgIleLeuProAsnLeuIlyLeuGlyCysGlnArgAspValLeuAsnLysAspLeuMet 180
QY 541 CAGCAGAAATGGAGTTGTTATGTTAAATGCCAGCTATACCTGCTCAAGAGCTGACTTT 600
Db 181 GlnGlnAsnGlyIleGlyTyraValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGATCTCATTTCTCGCGTGGCTGCTGTAATGACAGCTTTGTGAGAAATTTTG 660
Db 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluTyraIleLeu 220
QY 661 CCGGTGTTGGCAAAATCAGTATGATTTCATGAGAAACAAAGCTCCATATGATGTTG 720

Db 221 ProThrLeuAspLysSerValAspPheIleGluTyraValAlaValAspAsnGlyCysVal 240
QY 721 CTATGTCACCTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCTACATATG 780
Db 241 LeuIleHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 260
QY 781 AAGAGATGAGCATGCTTATGATGAAAGCTTACAGATTTGTGAAGAAAAAGACTACT 840
Db 261 LysArgMetCaspMetSerLeuAspGluAlaTyraArgPheValLysGluLysAspProThr 280
QY 841 ATATCTCAAACTTCAATTTTCTGGGCCAATCTCGACTGACTATGAGAAAGATTAAAGAC 900
Db 281 IleSerProAsnPheAsnPheMetGlyGlnLeuMetAspTyraGlySerThrIleAsnAsn 300
QY 901 CACACTGAGCATCAGGCGCAAGAAAGCAACTCAAGCTGCTGCACTCGAGAAAGCAAT 960
Db 301 GlnThrGlyMetSerGlyProLysSerTyraLeuTyraLeuMetHisLeuAspLysProSer 320
QY 961 GAACGTGCTGCTGCTGCTCAGAGGGTGGCAGAAAGAGAGAGCGCCCTCAGTCCACCC 1020
Db 321 GluProValProAlaAspSerGluGlyTyraPlySerAlaLeuSerLeuSerProPro 340
QY 1021 TGTGCGACTCTGCTACTCTCAGAGGAGCAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1080
Db 341 CysAlaAsnSer---ThrSerGluAlaSerGlyGlnArgLeuValHisProAla----- 357
QY 1081 CCGAGCGTCCAGCGCTGACAGCCGTGCTGTTAGAGACAGCCCGCTGTAACAGCGCTC 1140
Db 358 ---SerValProSerLeuGlnProSerLeuGluAspSerThrTyraLeuLysArgSerPhe 376
QY 1141 AGTGGCTGCACTGTCGCGCAGACAGCGTGGAGAGCAATAGCTTCMAAGCGTTCCTTC 1200
Db 377 SerGlyLeuGlnLeuSerSerGlyLysLeuGluAspSerThrTyraLeuLysArgSerPhe 396
QY 1201 TCTCTGAAATCAAACTCAGTTTATTCATTCAGCCAGATGGCAGCATCTTACATGCTTC 1260
Db 397 SerLeuAspIleLysSerValSerTyraSerAlaSerMetAlaIleSerLeuHisGlyPhe 416
QY 1261 TCTCATCAGAGATGCTTTGGAATCACTCAAACTTCCATCTGATCGATGGAGCAAC 1320
Db 417 ---SerSerGluGlnIleAlaIleAspTyraCysLysProSerAlaThrLeuAspGlyThrAsn 435
QY 1321 AAGCTATGCCAGTTTCCCTGTTTACGAAACTATCGAGAGAGCTCCGGAACAGTCC 1380
Db 436 LysLeuCysGlnPheSerProValGlnGluValSerGlnGlnSerProGluThrSerPro 455
QY 1381 GATTAAGAGAGAGCCAGCATCCCAAGAAAGCTGAGACCGCCAGGCTTCAGACAGCAG 1440
Db 456 AspLysGluGlnAlaHisIleProLysGlnProGlnProProArgProSerGluSerGln 475
QY 1441 AGCAAGCATTTGATTCGCTCAGAACCAAGAGTGGCAGCCGCCAGAGTCCCTTTTA 1500
Db 476 ValThrArgLeuHisSerValArgThrGlySerSerGlySerThrGlnArgProPhePhe 495
QY 1501 TCTTCACCTGATGGAAGTGGAGCGTGGAGACAAATTAACAACAGCTTCTTTCCG-- 1559
Db 496 SerProLeuHisArgSerGlySerValGluAspAsnIlyHisThrAsnPheLeuPheGly 515
QY 1560 CTTTCACACGAGCAGCAGACCTCAAGAGTGTGCTGCGCTGAGGCTTTAAGGGCTGGCA 1619
Db 516 ProPheHisGlnProAlaThrProHisGlnValCys-ArgAlaIleThrLeuLysGlyTrpHis 535
QY 1620 CTGGAATATCTGGCCCCCAGACCTTCACCTTCTCTGACCAAGAGCTGATTTTGC 1679
Db 535 AspArgIleLeuAlaProGlnSerSerAlaProSerLeuThrSerSerTrpTyraPheAl 555
QY 1680 CACAGAGTCTCTACATTCATCTGCTCAGCACTCAAGCACTCAAGAGAGAGTCCAGTACTC 1739
Db 555 atmGluProSerHisLeuIlySerAlaSerAlaIleTyraGlyGlyAsnSerSerTyraSe 575
QY 1740 TGCCTACAGCTGACGAGCAGCTGCCACTTGGCGAGACCAAGTTATTTCTGCGCAGGCG 1799

[illegible]

OC	Mammalia;Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
KN	SEQUENCE FROM N.A.
RP	STRAIN=BALB/C;
RC	Matsumuchi T., Musikacharen T., Johnson T.R., Kraft A.S.,
RA	Yoshikai Y.;
RT	"A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT	Activation in Macrophages.";
RL	Mol. Cell. Biol. 20:699-700(2001).
DR	EMBL; AF345853; AK35054.1; ..
DR	HSSP; Q16828; IMKP.
DR	MGI; MGI:1917936; Dusp16.
DR	GO; GO:0005737; C:cytoplasm; IDA.
DR	GO; GO:0005634; C:nucleus; IDA.
DR	GO; GO:0005515; P:protein binding; IDP.
DR	GO; GO:0000188; P:inactivation of MAPK; IDA.
DR	InterPro; IPR000340; DS phosphatase.
DR	InterPro; IPR008343; MAPK_phosph.
DR	InterPro; IPR001763; Rhodanese-like.
DR	InterPro; IPR000387; TYR_phosphatase.
DR	Pfam; PF00782; DSPC; 1.
DR	Pfam; PF00581; Rhodanese; 1.
DR	PRINTS; PR01764; MAPKPHPTASE.
DR	SMART; SM00195; DSPC; 1.
DR	SMART; SM00450; RHOD; 1.
DR	PROSITE; PSS0206; RHODANASE_3; 1.
DR	PROSITE; PSS00383; TYR_PHOSPHATASE_1; 1.
DR	PROSITE; PSS00056; TYR_PHOSPHATASE_2; 1.
DR	PROSITE; PSS00504; TYR_PHOSPHATASE_DUAL; 1.
DR	Hydrolase; Kinase.
QW	SEQUENCE 355 AA; 39502 MW; D5C29AE215CA285 CRC64; .

Alignment Scores:

Pred. No.:	3,586-103	Length:	355
Score:	1,455.50	Matches:	314
Percent Similarity:	49.02%	Conservative:	12
Best Local Similarity:	47.22%	Mismatches:	11
Query Match:	40.26%	Indels:	328
DB:	11	Gaps:	1

Oy	1	ATGGCCATGATGATGATGGAACCTCAAAATGCTTACTGAGAGTTGGTGGCTCTCGAA	60
Db	1	MetAlHhISgluMetIleGlyThrGlnIleValThrGluSerIleValAlaIleuLeuGlu	20
Oy	61	AGTGAACGGAAAAAGTGCCTCTTAATGATAGCCGGCCATTGTGGAATACATACATCC	120
Db	21	SeTcGlyThrGluIlyValIleuLeuIleAspSerArgProPheValGlyIlyAsnThrSer	40
Oy	121	CACATTTTGGAAAGCCATTATATCTCACTGCTCCAAAGCTTATGAAAGCGAAGTTGCACAG	180
Db	41	HhISleuGluAlaIleAsnIleAsnIcySerIlyIleuMetIlyAspArgIleuGlnGln	60
Oy	181	GAAAGATGTATTAACAGAGCTATCCAGCACTTACGGAAACATAAGTTGATGATTCAT	240
Db	61	AspArgValIleuIleThrGluIleuIleGlnHisSerAlaIlyHisIlyValAspIleAsp	80
Oy	241	TGAGTCAACAAAGGTGTGATTTACATCAAAAGCTCCCAAGATGTGTGCTCTCTCTTCA	300
Db	81	CyAsnGlnIlyValIleValIlyThrAspGlnSerSerGlnAspValGlySerIleuSerSer	100
Oy	301	GACTGTTTTCTCACTGTACTTTCTGGGTAACTGCGAAGAGCTTCAACTCTGTTCACCTG	360
Db	101	AspCyAspPheIleuThrValIleuIleuIlyIlyLeuGluIuArgSerPheAsnSerValHisIleu	120
Oy	361	CTTGCAAGGTGGGTTGCTGAGTTCTCTGTTGTTTCCCTGGGCTCTGTGAAAGGAAAAATCC	420
Db	121	LeuAlaGlyIlyPheAlaGluIupheserArgCyAspIleProGlyIleuCyGluGlyIlySer	140
Oy	421	ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTTGCAACATTGGGCGAACCC	480

Db 141 ThrLeuValProthrCysIleSerGlnProCysLeuProValAlaIleuIleGlyProthr 160
 QY 481 CGAATCTTCCCAATCTTTATCTTGGCTGCCAGCAGATGCTCCCAACAGAGCTGATA 540
 Db 161 ArgIleLeuProAsnLeuIleuGlyCysGlnArgValLeuAsnLeuSapLeuMet 180
 QY 541 CAGCGAATGGGATTGGTTAATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTT 600
 Db 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProIleProAspHe 200
 QY 601 ATCCCGAGCTCATTTCTCCCTGCTGCTGTGAATAGCAGCTTTGTGAGAAATTTTG 660
 Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIleLeu 220
 QY 661 CCGTGGTTGACAAATACAGATTTCATTGAGAAAGCAAAAGCCCTCCCAATGATGTGT 720
 Db 221 ProThrLeuAspIleSerValAspPheIleGlyValAlaIleAlaIleTyrIleMet 240
 QY 721 CTATGACACTGTAGTGGATCTCCGCTCCGACCATGCTATGCTCATCATCATG 780
 Db 241 LeuIleHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleTyrIleMet 260
 QY 781 AAGAGGATGGACATGTCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAAGACTACT 840
 Db 261 LysArgMetAspMetSerLeuAspGlnAlaTyr----- 271
 QY 841 ATATCTCAAACTTCAATTTCTGGCCCACTCTGACATATGAGAAAGATTAAAGAC 900
 Db 271 ----- 271
 QY 901 CAGACTGAGCATCAGGGCCAAAGACAAACTCAGCTGTGACCTGAGAGAACCAAT 960
 Db 271 ----- 271
 QY 961 GAACCTGTCCGTGTGTCTCAGAGGGTGAACAGAAAAAGCAGAGCCCTCAGTCCACC 1020
 Db 271 ----- 271
 QY 1021 TGTGCCACTCTGCTACTCCTCAGAGGACAGAGCAAAAGCCGTGATCCGCCACGCTG 1080
 Db 271 ----- 271
 QY 1081 CCCAGCGTGCACAGCTGAGCCGTGTGTAGAGAGAACGCCCGTGTGACAGCGCTC 1140
 Db 271 ----- 271
 QY 1141 AGTGCGTGCACCTGTCCGACAGAGCTGGAAGACGACATTAAGCTCAAGCTTCTTC 1200
 Db 271 ----- 271
 QY 1201 TCTCTGATATCAATCAGTTTCATATTCAGCCACGATGACATCTTACATGCTTC 1260
 Db 271 ----- 271
 QY 1261 TCCTCATCAGAAAGATGCTTTGGAATATCAAACTTCACTACTGTGATGGAGCAAC 1320
 Db 271 ----- 271
 QY 1321 AAGCTATGCCAGTTCTCCCTGTTCAGGAAGTATGAGAGACACTCCCGAAACAGTCT 1380
 Db 271 ----- 271
 QY 1381 GATAGAGAGAAAGCAGCATCCCAAGACTGACAGCCGACGCTTCAGACAGCCAG 1440
 Db 271 ----- 271
 QY 1441 AGCAAGCATTTGATTTGGATCAGAACACAGACAGATGAGACCGCCCAAGATGCTTTT 1500
 Db 271 ----- 271
 QY 1501 TCTCAGTGCATGGAAGTGGAGCGTGGAGACAAATTACACACAGCTTCTTTTGGC 1560
 Db 271 ----- 271

QY 1561 CTTTCCACAGCCAGCAGACACTTACAGAAATCTGTGCTGGCTTGAAGGCTGGCAC 1620
 Db 271 ----- 271
 QY 1621 TCGAATATTTGAGCCGCCACAGACTCTACCCCTTCCCGACACAGACTGTATTTTGC 1680
 Db 271 ----- 271
 QY 1681 ACAGAGTCTCAGACTTCTACTCTGCTCAGACCATTTACGAGAGCAGTGCAGTTACTCT 1740
 Db 271 ----- 271
 QY 1741 GCTTACAGCTCAGCCAGCTGCCCATTTGGAGACCAAGTCTATTTCTGTGGCAGCGG 1800
 Db 272 ----- 272
 QY 1801 CAGAACCAAGTACAGAGCTGACTGCGCGGAGCTGGCATGAAGAGCCCTTTGAA 1860
 Db 274 GlnIleProthrAspAlaAspSerArgArgThr-GlyMetIleSArgAlaProLeuIle 1920
 QY 1861 AAGCATTTTAAAGCAGAACTGCCAATGAAATTTGGAGAGCATCATGTCAGAGAC 1920
 Db 293 SSerSerLeuAsnAlaGlyAlaAlaIleTyrAsnLeuGlnArgAlaLeuCysArgArg 313
 QY 1921 AGGTACGGGAGAGCTGGGAAAGTGGGCACTCAGTCTACCTTTTGGGCGACATGAA 1980
 Db 313 GlnIleProGlyArgSerTyrAlaArgTyrAlaAlaSerProAlaSerProAlaIleTyr 333
 QY 1981 ATCATTAGAGTCT 1993
 Db 333 gSerSerArgSer 337

RESULT 8

086SS8

ID 086SS8

AC 086SS8

DT 01-JUN-2003

DT 01-JUN-2003

DT 01-OCT-2003

OS Homo sapiens

OC Eukaryota

OC Mammalia

OC NCI_Taxid=9606

RP SEQUENCE FROM N.A.

RC Tissue=Brain, and Astrocytoma

RA Strauberg R.

RL Submitted (JAN-2003)

DR EMBL; BC045110; AAH45110.1; -

DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.

DR InterPro; IPR000340; DS phosphatase.

DR InterPro; IPR008343; MARK phosph.

DR InterPro; IPR002965; P-rich exten.

DR InterPro; IPR001763; Rhodanese-like.

DR Pfam; PF00782; DSPC; 1.

DR PRINTS; PRO1764; MARKPHPTASE.

DR SMART; SM00195; DSPC; 1.

DR SMART; SM00450; RHOD; 1.

DR PROSITE; PS50206; RHODANES_3; 1.

DR PROSITE; PS00383; TYR PHOSPHATASE_1; 1.

DR PROSITE; PS50056; TYR PHOSPHATASE_2; 1.

DR PROSITE; PS50054; TYR PHOSPHATASE DUAL; 1.

SQ SEQUENCE 625 AA; 65826 MW; C7C808407B724PFC CRC64;

Alignment Scores:

Ptd. No.: 3,13e-93

Score: 1327.00

Percent Similarity: 58.54%

Length: 625

Matches: 314

Conservative: 94


```

QY 1330 CAGTTCTCCCTGTCAGAACTATCGAGACAGCTCCGAAAACAGTCTGATGAAGAG 1389
DB 438 -----ProSerProSerProSerPro 445
QY 1390 GAAGCAGCATCCCAAG-----AGCTGCAGACCCGACGCTTCAGACAGCAG 1440
DB 446 Asp---SerValProGluCybArgProArgProArgArgArgProProAlaSerSer 464
QY 1441 AGCAAGCGA-----TTGCATTGGGTGAGAACCAAGCAGCTGACCCGACAGAGTCC 1494
DB 465 ProAlaArgSerProAlaArgLeuGlyLeuAlaSerProAlaThrAlaArgGln--- 483
QY 1495 CTTTATCTCAGCATCGATCGAAGTGGAGGTGGAGACAATTACACACAGCTTCCTT 1554
DB 484 -----ThrProArgHis----- 487
QY 1555 TTGGGCTTTTCACACAGCAGCAGACCTTCAGAGTCTGCGCTG----- 1602
DB 488 ---GlyLeuSerAla-----LeuSerAlaProGlyLeuProGlyProGly 501
QY 1603 -----GGCTTAAAGGCTGCGACTCGGATATCTGGCCCCGACAGCTTACCCCT 1653
DB 502 GlnProAlaGlyProGlyGlyTyrValProProLeuAlaSerPro-----GlyThrPro 519
QY 1654 TCCCTGACGACGAGCTGTATTTTGGCAGACAGAGCTTCACAC-----TTCTAC 1701
DB 520 SerProAlaGlyProTyrCybSerSerProGluGlyAlaGlnGlyProGlyAlaValPhe 539
QY 1702 TCTGCC-----TCAGCCATCTACGGA 1722
DB 540 SerAlaPheGlyArgValSerAlaGlyAlaProGlyProGlyAlaSerSerSerGly 559
QY 1723 GGC----- 1725
DB 560 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 579
QY 1726 -----AGTGCAGATTACTCTGCTACAGCTGCGACGCGCCACTTGGGAGAC 1776
DB 580 SerAlaSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 599
QY 1777 CAACTATTCTGTGCGCAGCGCGCAGAACCAAGTGACAGAGCTGACGCGCGAGC 1836
DB 600 SerSerSerAlaPheArgArg-----AspValArgThrGly 612
QY 1837 TGGCAGTAAGAGAGCCCCCTTTGAAAAGCAGTTTAAACCGCAAGCTGCCAAATGAATTT 1896
DB 613 TrpProGluGluProAlaAlaAspAlaGlnPheGlyArgSerCybGlnMetGluPhe 632
QY 1897 GGAGAGAGCATCATGTGAGAGAACAGGTCAAGG---GAAGAGCTGGGGAAGGGCGAGT 1953
DB 633 GlnGlnGly---MetValGlnGlyArgAlaArgGlyGlnGlnAlaAlaLeuGlyLys 651
QY 1954 CAGTCTAGCTTTTGGGCGACATGGAATCATTTGAGGCTTCC 1995
DB 652 GlnThrSerPheSerGlySerValGlnValIleGlnValSer 665

```

```

RA Yoshikai Y.:
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages."
RL Mol. Cell. Biol. 20:6999-7009 (2001).
DR EMBL, AF345954; AAK35055.1; -.
DR HSSP, Q16828; IMKP.
DR MGI, MGI:1917936; Dusp16.
DR GO, GO:0005737; Cytoplasm; IDA.
DR GO, GO:0005634; Cytosol; IDA.
DR GO, GO:000515; Protein binding; IPT.
DR GO, GO:000188; P:inactivation of MAPK; IDA.
DR InterPro, IPR000340; DS phosphatase.
DR InterPro, IPR008343; MAPK phosph.
DR InterPro, IPR001763; Rhodanese-like.
DR InterPro, IPR000387; Tyr phosphatase.
DR Pfam, PF00782; DSPC; 1.
DR Pfam, PF00581; Rhodanese; 1.
DR PRINTS, PR01764; MAPKPHPTASE.
DR SMART, SM00195; DSPC; 1.
DR SMART, SM00450; RHOD; 1.
DR PROSITE, PS50206; RHODANSE_3; 1.
DR PROSITE, PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE, PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE, PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydrolase; Kinase.
SQ SEQUENCE 300 AA; 33624 MW; 4C61846ACDF0F456 CRC64;

Alignment Scores:
Pred. No. 5,79e-77 Length: 300
Score: 1114.00 Matches: 259
Percent Similarity: 40.75% Conservative: 12
Best Local Similarity: 38.95% Mismatches: 11
Query Match: 30.82% Indels: 383
Gaps: 2

US-10-029-345A-108_COPY_538_2532 (1-1995) x Q9AG15 (1-300)
QY 1 ATGGCCCATGAGATGATGATGGAAGTCAATTTGTTACTGAGAGGTGTGGCTGTGGA 60
DB 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnSerLeuValAlaLeuGln 20
QY 61 AGTGAACGGAAGGATGCTGCTAATTGATGAGCGGCAATTTGTGGAATACATATCC 120
DB 21 SerGlyThrGlnIleValLeuLeuIleAspSerArgProPheAlaGlnIleThrSer 40
QY 121 CACATTTGGAAGCCATTAATATCAATGCTCCCAAGCTTATGAGGAGGTTGCAACG 180
DB 41 HisIleLeuGlnIleAlaIleAsnIleAsnCyserIleLeuMetIleArgIleGln 60
QY 181 GAAAGGTGTAATTCAGAGCTCATCCAGCATTCAGCGAACAATAGGTTGACATTTGAT 240
DB 61 AspArgValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisIleValAlaPhe 80
QY 241 TGAAGTCAGAAAGGTTAGTTAGATCAAGTCAAGCTCCCAAGATGTTCCTCTCTCA 300
DB 81 CysAlaGlnIleValValValIleTyrAspGlnSerSerGlnAspValGlySerLeuSer 100
QY 301 GACTGTTTCTCATCTGACTCTGGGTAACTGGAGAGAGCTTCAACTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnArgSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGAGGTGGTTCGATGATTCCTGCTGTTTCCCTGCGCTCTGTAGAAGAAAATC 420
DB 121 LeuAla----- 122
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCTTGCTTCTACTGTTGCCAATTTGGGCAACC 480
DB 122 ----- 122
QY 481 CGAATCTTCCCACTTATCTTGGCTGCCAGCGAGATGCTCTCAACAGAGCTGATN 540
DB 123 -----AspLeuMet 125

```


Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnSerLeuValAlaLeuGlu 20
QY 61 AGTGAACCGAAAAAGCTGCTTAATGATAGCCGGCATTGTGGAAATACATACATCC 120
Db 21 SerGlyThrGlnValLeuLeuIleAspSerArgProPheValGlnGlyrRanThrSer 40
QY 121 CACATTTGGAGCCATTAAATATCAATGCTCCCAAGTTTGAAGGAGGTTGCAACAG 180
Db 41 HisIleLeuGlnAlaIleValLeuIleAsnGlySerLeuMetCysArgArgLeuGlnGln 60
QY 181 GACAAAGCTGTAATTCAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
Db 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisLysValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATGATTACGATCAAGCTCCCAAGATGTTGGCTCTCTCTTCA 300
Db 81 CysAsnGlnLysValValValValValAspGlnSerSerGlnAspValGlySerLeuSerSer 100
QY 301 GACTGTTTCTCACTGACTCTGGGTTAACTGGAGAGAGCTTCAACTGTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnArgSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGTGGGTTGCTGAGTTCTCTGTTGTTTCCCTGCTGTGAGGAAATCC 420
Db 121 LeuAlaGlyLysPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlnGlyLysSer 140
QY 421 ACTCTAGTCCCTACCTGATTTCTGAGCTTGTCTTACTGTTGCAACATTTGGCCCAAC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTTCCCAATCTTTATCTTGCTGCGCAGCAGAGATGCTCCCAAGAG 531
Db 161 ArgIleLeuProAsnLeuTrpLeuGlyCysGlnArgAspValLeuAsnLys 177

RESULT 12

ID 0960S2 PRELIMINARY; PRT: 143 AA.

AC 0960S2; 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Truncated MAPK phosphatase 7.
GN MKP7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Montpécit A., Boily G., Simmet D.;
RT "A detailed transcriptional map of the chromosome 12p12 tumor
suppressor locus."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY038927; AAK69770.1; -
DR GO; GO:0017017; F-MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P-protein amino acid dephosphorylation; IEA.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00450; RHOD.1.
DR PROSITE; PSS0206; RHODANASE_3; 1.
SQ SEQUENCE 143 AA; 16077 MW; 5213A213AA7E5974 CRC64;

Alignment Scores:

Pred. No.: 1,2e-41 Length: 143
Score: 653.00 Matches: 140
Percent Similarity: 70.35% Conservatve: 0
Best Local Similarity: 70.35% Mismatches: 3
Query Match: 18.06% Indels: 56
DB: 4 Gaps: 1

US-10-029-345A-108_COPY_538_2532 (1-1995) X 0960S2 (1-143)

QY 1 ATGGCCATGAGATGATGGAACTCAATTTGTTACTGAGAGGTGGGCTGTGCGAA 60
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnSerLeuValAlaLeuGlu 20
QY 61 AGTGAACCGAAAAAGCTGCTTAATGATAGCCGGCATTGTGGAAATACATACATCC 120
Db 21 SerGlyThrGlnValLeuLeuIleAspSerArgProPheValGlnGlyrRanThrSer 40
QY 121 CACATTTGGAGCCATTAAATATCAATGCTCCCAAGTTTGAAGGAGGTTGCAACAG 180
Db 41 HisIleLeuGlnAlaIleValLeuIleAsnGlySerLeuMetCysArgArgLeuGlnGln 60
QY 181 GACAAAGCTGTAATTCAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
Db 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisLysValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATGATTACGATCAAGCTCCCAAGATGTTGGCTCTCTCTTCA 300
Db 81 CysSerGlnLysValValValValValAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTCTCACTGACTCTGGGTTAACTGGAGAGAGCTTCAACTGTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnArgSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGTGGGTTGCTGAGTTCTCTGTTGTTTCCCTGCTGTGAGGAAATCC 420
Db 121 LeuAlaGly--Ala----- 124
QY 421 ACTCTAGTCCCTACCTGATTTCTGAGCTTGTCTTACTGTTGCAACATTTGGCCCAAC 480
Db 124 ----- 124
QY 481 CGAATTTTCCCAATCTTTATCTTGCTGCGCAGCAGAGATGCTCCCAAGAGCTGAT 540
Db 125 -----AspA 126

RESULT 13

ID 08ST19 PRELIMINARY; PRT: 657 AA.

AC 08ST19; 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F08B1.1a.
GN F08B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Br1etol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans a platform for
investigating biology". The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Br1etol N2;
RA Chissoe S.;
RT "The sequence of C. elegans cosmid F08B1.1";
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Br1etol N2;
RA Waterston R.;
RT "Direct Submission.";


```

Db      371  ThrmctaspleugllyleuprohisatpProlyslaleucllyeuproserarglegly 390
Oy      1417  ACCGCC-----AGCCTTCAGACAGC---CAGAGCAAGGATTTGATTCGGTCGA 1464
Db      391  ThrservalalagluleuproserProserThrluenserargleuserhe-AsnG1 410
Oy      1465  ACCAGCAGCAGTGCGCACCGCCGACAGAGTCCCTTTATCTCCACTGCATGCAAGTGAGAC 1524
Db      410  yProglualalialalaproserthr---Proilleu----- 421
Oy      1525  GTGAGAGCAATTACACACACAGCTTCCTTTTCGGCTTTCACAGCAGACAGACCTC 1554
Db      422  -----AsmpherthAsmProCyasphAsmSerProilleProvalalaserse 439
Oy      1585  ACAGAGTCTGCTGCTGGGCGCTTAAGGCTGACCTGCATATCTTGGCCCCAGACC 1644
Db      439  rAsgluValille----- 443
Oy      1645  TCTACCCCTTCCTGACACGACGCTGATTTTGGCAGAGATCCTGACCTTACTCT 1704
Db      444  -LeuthrleuProthrProalalala-----SerseSerseSerthrse 458
Oy      1705  GCCTCAGCCA-----TCTACGAGGACGAGTGCAGT 1734
Db      458  rsergluProserPheaspheserSerPhegluSerSerSerSerSerleValVa 478
Oy      1735  TACTTGCTTACAGCTGACGACGACCTGCCACTTGGCAGACCAAGCTTATCTGCGCC 1794
Db      478  lgluanProPhePhealaserThrgluValProalglySerSerSerleSerthrPr 498
Oy      1795  AGCGCGAGAGCAAGGACAGAGCTGACTCGCGCGAGCGAGCTGCATGAGAGAGCCCC 1854
Db      498  oserdly---serlinserrhrProalaserlaserSerSerlaserArgCysAr 517
Oy      1855  TTTGAAA 1861
Db      517  gmecllys 519

```

RESULT 15

```

O7SZF3  PRELIMINARY;      PRT;      367 AA.
AC  O7SZF3;
DT  01-OCT-2003 (TrEMBLrel. 25, Created)
DT  01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Hypothetical protein.
OS  Brachydanio rerio (Zebrafish) (Danio rerio).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC  Cyprinidae; Danio.
OX  NCBI_TaxId=7955;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=AB; TISSUE=Body;
RX  MEDLINE=2388257; PubMed=12477932.
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieff F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallaby S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Falley J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Bouffard G.G.,
RA  Rodriguez A.C., Gilmour J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA  Kryzhanetskiy M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA  Jones S.J., Maitra M.A.;

```

```

RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN      [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=AB; TISSUE=Body;
RA  Strausberg R.;
RL  Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL, BC052477; AAH52477.1; -.
KW  Hypothetical protein.
SQ  SEQUENCE 367 AA; 40538 MW; F8997A415DAD6AC CRC64;

```

Alignment Scores:

Pred. No.:	8,11e-29	Length:	367
Score:	487.50	Matches:	114
Percent Similarity:	53.82%	Conservative:	69
Best Local Similarity:	33.53%	Mismatches:	124
Query Match:	13.49%	Indels:	33
DB:	13	Gaps:	8

US-10-029-345A-108_COPY_538_2532 (1-1995) x O7SZF3 (1-367)

```

Oy      43  TTGGTGCTCTGCTGAAAAGTGGAAAGGAAAGTCTGTAATTGATAGCGGCATT 102
Db      15  LeuysargleuMetlysaAspArglyAlaLysCysleuLeuAspCysArgSerPhe 34
Oy      103  GTGGAATPACATATCATCCCATTTTGGCAAGCATTATATACCTGCTCAAGCTTATG 162
Db      35  LeuAlaPheSerIleaglYhisleuArgglYalaValanIleArgCysAserThrIleVal 54
Oy      163  AACCGAAG-----TTGCACAGACGACAAAGTTTATATACAGAGCTATCAG 210
Db      55  ArgArgArgAlaLysglYserValSerleuAspGlnIleleuSerGlyAsp----- 71
Oy      211  CATTCAGCAAAATPAGAGTTGACATTGATTCAGCTGAGAAAGTTGATTCAGTCA 270
Db      72  AspGluAlaArgSerArgleuSerleuYserIleValIleleuYArgGlu 91
Oy      271  AGCTCCCAAGATGTTGCTCTCTTCAGAGCTGTTTCTGCTGATTCCTGAGTAA 330
Db      92  ArgSerSerAspThrAsnThrmetylaAspSerThrIleThrleuValHisAsnAla 111
Oy      331  CTG---GAGAAGAGGTTCAACTC---CTTCACCTGCTTCAGAGTGGTTTGGTGAATC 384
Db      112  LeuysArgAspThrPheSerThrgluValYleuLeuysglYglYTrAspArgPhe 131
Oy      385  TCTGCTGTTTCTCGTGCCTCTGT-----GAGGAAAA 417
Db      132  SerThrgluTrProAspTrYrCysleuYThrArgThleuSerValSerSerGln 151
Oy      418  TCACTCTAGTCCTTACCTGATCTTCTCAGACCTTTCCTTTCCTGTTCCAAATTCGGGCA 477
Db      152  SerSerMetGluSerSerCysleuSerCysAlaThrProGlnHisAspGlnolYlYPro 171
Oy      478  ACCGAATTTCTCCCAATCTTTATCTTGGCTGCGCAGCGAGATGTCCTGCAAGAGAGCTG 537
Db      172  ValgluIleleuProPheleuPheleuGlySerAlaLeuHisAlaSerLysAlaSerPhe 191
Oy      538  ATCAGAGAGAAATGGATTTGTTGTTGTTAAATGCCAGCTATACCTTCGCAAAAGCTGAC 597
Db      192  LeuAspArgMetCylLleSerAlaLeuLeuSerValSerSerAsnCyProAsn---His 210
Oy      598  TTTATCCCGAGTCTCATTTCTCGTGCCTGCGCTGCGTGAATGACAGCTTTTGGAAAAATT 657
Db      211  PheGluGluAspTrYrClnTrYrCysIleleProValGluAspAsnHisLysGluAspIle 230
Oy      658  TTGCGGTGTTTGAACAAATCAGTACATTCATTTGAAGAAAGCAAGGCTCAATGATGT 717
Db      231  SerSerThrPheIleGluAlaIleGluPheIleAspSerValLysAspSerAsnGlyArg 250
Oy      718  GTTCTAGTCACTGTTTATGCTGAGATCTCCGCTCGCGCCGCAACATCGTATCGCTTATC 777
Db      251  ValLeuValHisCysGlnAlaGlyLleSerArgSerAlaThrIleCysleuAlaTrYleu 270

```

```
OY 778 ATGAGAGATGACATGCTTTAGATGAGCTTACAGATTGTGAAAGAAAAAGACT 837
Db 271 MetLysIleValArgValArgLeuGluAlaPheGluValIleValArgArgSer 290
OY 838 ACTATATCTCCAACTTCAATTTCTGGGCCCACTCCTGAGACTATGAGAGAAGATTAAAG 897
Db 291 IleIleSerProAsnPheserPheMetGlyGlnLeuGlnPheGlnSerGlnValLeu 310
OY 898 AACCAAGCTGAGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGACCTGAGAGAGCCA 957
Db 311 ---AlaThrSerCysSerValGluAlaIleSer----- 320
OY 958 AATGAACCTGTCCCTGCTGTCTCAGAGGCTGACAGAGAAAGCGAGAGCGCCCTCAGTCCA 1017
Db 321 -----ProSerAlaSerLeuGlyProLysSerSerSerThrProIleSerPro 336
```

Search completed: June 21, 2004, 12:40:18
Job time : 138.154 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:40:31 ; Search time 94.9018 Seconds

(without alignments)
11869.448 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532
Perfect score: 3615
Sequence: 1 atggcccaatgagatgattg9.....tggaatcatctgagctctcc 1995

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 2327084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_epool_p/US10029345/runat_21062004_122818_4331/app_query.fasta_1.3278
-DB=Published Applications AA -QFMT=fasta -SUFFIX=rabp -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human4.cdt -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10029345 @CGN_1_1_21 @runat_21062004_122818_4331
-NCPU=6 -ICPU=3 -NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

Published Applications_AA:*

- 1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/prodata/1/pubppa/PCR_NEW_PUB.pep:*
- 3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/prodata/1/pubppa/PCRUS_PUBCOMB.pep:*
- 7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

1	3406	94.2	665	9	US-09-816-494-2	Sequence 2, App1
2	3406	94.2	665	12	US-09-964-217-2	Sequence 2, App1
3	3406	94.2	665	12	US-10-072-012-680	Sequence 680, App
4	3406	94.2	665	12	US-10-168-506-14	Sequence 14, App1
5	3406	94.2	665	12	US-10-343-357-7	Sequence 7, App1
6	3406	94.2	665	15	US-10-377-072-26	Sequence 26, App1
7	3406	94.2	665	16	US-10-257-026-2	Sequence 2, App1
8	3406	94.2	665	16	US-10-648-593-247	Sequence 240, App
9	3406	94.2	665	16	US-10-648-593-247	Sequence 247, App
10	3406	94.2	665	12	US-10-072-012-679	Sequence 679, App
11	3406	94.2	665	12	US-10-072-012-703	Sequence 703, App
12	3406	94.2	665	12	US-10-425-114-54204	Sequence 54204, A
13	3399	94.0	665	12	US-10-072-012-681	Sequence 681, App
14	3399	94.0	665	15	US-10-094-749-2312	Sequence 2312, Ap
15	3379.5	93.5	662	12	US-10-072-012-258	Sequence 258, App
16	3358.5	92.9	660	12	US-10-072-012-256	Sequence 256, App
17	3322	91.9	672	12	US-10-296-115-1259	Sequence 1259, Ap
18	3079.5	85.2	660	12	US-10-072-012-682	Sequence 682, App
19	2954.5	81.7	677	12	US-10-072-012-683	Sequence 683, App
20	2594	71.8	517	9	US-09-964-277-21	Sequence 21, App1
21	1326	36.7	625	12	US-10-072-012-639	Sequence 639, App
22	1302	36.0	663	12	US-10-072-012-700	Sequence 700, App
23	1297	35.9	253	15	US-10-108-260A-4872	Sequence 4872, Ap
24	1103.5	30.5	616	12	US-10-072-012-266	Sequence 266, App
25	917	25.4	501	12	US-10-072-012-702	Sequence 702, App
26	807	22.3	155	9	US-09-964-277-7	Sequence 7, App1
27	735.5	20.3	461	12	US-10-072-012-701	Sequence 701, App
28	654.5	18.1	169	14	US-10-346-356-15	Sequence 15, App1
29	654.5	18.1	170	9	US-09-775-925-26	Sequence 26, App1
30	654.5	18.1	170	9	US-09-847-519A-11	Sequence 11, App1
31	654.5	18.1	170	12	US-10-655-073-16	Sequence 16, App1
32	654.5	18.1	170	14	US-10-314-058-14	Sequence 14, App1
33	654.5	18.1	170	14	US-10-405-808-16	Sequence 16, App1
34	619	17.1	155	9	US-09-964-277-6	Sequence 6, App1
35	619	17.1	155	9	US-09-955-732-6	Sequence 6, App1
36	469	13.0	444	12	US-09-964-899-47	Sequence 47, App1
37	469	13.0	442	9	US-10-058-270A-130	Sequence 130, App
38	469	13.0	482	14	US-10-346-356-2	Sequence 2, App1
39	469	13.0	482	16	US-10-648-593-164	Sequence 164, App
40	466.5	12.9	381	14	US-10-184-832-2	Sequence 2, App1
41	455	12.6	394	9	US-09-736-457-805	Sequence 805, App
42	455	12.6	394	9	US-09-902-941-805	Sequence 805, App
43	455	12.6	394	9	US-09-849-626-805	Sequence 805, App
44	455	12.6	394	12	US-10-283-017-805	Sequence 805, App
45	455	12.6	394	14	US-10-017-754-805	Sequence 805, App

ALIGNMENTS

RESULT 1
US-09-816-494-2
Sequence 2, Application US/09816494
Patent No. US20020034807A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OF INVENTION: 38592 AND 21117, NOVEL DUAL SPECIFICITY
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,858
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-09-816-494-2

Alignment Scores:
Pred. No.: 5.84e-234 Length: 665
Score: 3406.00 Matches: 663

Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 94.22%

Conservative: 1
Mismatches: 1
Indels: 0
Gaps: 0

US-10-029-345a-108_copy_538_2532 (1-1995) x US-09-816-494-2 (1-665)

```
OY 1 ATGGCCCATGAGATGATTTGGAACCAATTGTTACTGAGAGAGTTGGCTGCTGGGAA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
OY 61 AGTGAACGAAAAAGTGGCTTAATTTGATAGCCGGCAATTTGGGAATACAAATCATCC 120
DB 21 SerIleThrGluLysValIleLeuIleAspSerArgProPheValGluIleThrSer 40
OY 121 CACATTTGGAGCGCATTTATATCACTGCTCCAGCTTATGACGGAAGTTGCAACAG 180
DB 41 HisIleLeuGlnIleAlaIleAsnIleAsnIleCysSerLysLeuMetLysArgArgLeuGln 60
OY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGATTGACGAAACATAAGTTGACATTGAT 240
DB 61 AspLysValIleLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspLys 80
OY 241 TGCAGTCAGAAAGGTGTAGTATAGATCAAAAGCTCCCAAGATGTGCTCTCTCTCA 300
DB 81 CysSerGlnLysValValValValIleAspGlnSerSerGlnAspValAlaSerLeuSer 100
OY 301 GACTGTTTCTACGTAAGCTTCTGGGTAACTGAGAAAGAGCTTCACTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlnLysLeuGlnLysSerPheAsnSerValHisLeu 120
OY 361 CTTCAGAGTGGGTGTGCTGAGTCTCCGCTGTTTCCCTGCTCTGTGAAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlySer 140
OY 421 ACTGAGCTCTACCTGCAATTTCTCAGCCTTGTGCTTACTGTTGCAACATTTGGGCAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
OY 481 CGAATTCCTCCCAATCTTTATCTTGGCTGACGAGATGTCCTTCAACAGAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuLysLeuGlyCysGlnArgAspValLeuAsnLysGlnLeuMet 180
OY 541 CAGCAAAATGGATTGGTTATGTGTTAAATGCCAGCTATACCTGTCGAAAGCTGACATT 600
DB 181 GlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
OY 601 ATCCCGAGTCTCATTTCCGCGGTGCTGCTGATGACAGCTTTGTGAGAAATTTTGG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
OY 661 CCGTGTGGAACAATCAGTAGATTTCATTGGAAGCAAAAGCTCCCATGATGATGTT 720
DB 221 ProThrPheAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
OY 721 CTAGTGCACTGTTTGTGCTGGGATCTCCGCTCCGCAACATCGCTATCGCTATCATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIleMet 260
OY 781 AAGAGAGTGAACATGCTTTAGATGAAGCTTACAGATTGGAAGAAAAAGAACTACT 840
DB 261 LysArgMetAspMetSerLeuAspGlnAlaIleArgPheValLysGlnLysArgProThr 280
OY 841 ATATCTCCAACTTCATTTTCTGGGCAACTCTGGAATATGAGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspLysArgLysLysIleLysAsn 300
OY 901 CAGACTGAGGATATGAGGCGCAAGAGCAACTCAAGCTGCTGAGACCTGAGAAAGCAAT 960
DB 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlnLysProAsn 320
OY 961 GAACCTGTCTCTGTCTGAGAGGTTGACAGAAAGGAGAGCGCCCTCAGTCCACCC 1020
DB 1021 TGTGCGCACTGTGCTTACCTCAGAGCGACAGCAAGAAAGCCCGTGATCCCGCAGCGTG 1080
DB 341 CysAlaAspSerAlaThrSerGlnAlaIleGlnArgGlnArgProValHisProAlaSerVal 360
OY 1081 CCCAGCGTCCCGCAGCGTCCAGCGCTGCTGTTTAAAGAGACGCCCGGTGATCAGCGCTTC 1140
DB 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
OY 1141 AGTGGCTGCAACCTGTCGAGAGACAGCTGGAAGACAGCAATAAGCTCAAGCTTCTTC 1200
DB 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 400
OY 1201 TCTCTGATTCATTAATGATTTATATTTAGCCAGAGATGGACATCTTAATGAGCTTC 1260
DB 401 SerLeuAspLysSerValSerLysSerAlaSerMetAlaHisSerLeuHisGlyPhe 420
OY 1261 TCTCATCAGAAAGATGCTTGGAAATCTACAAACCTTCCACTACTGTGATGGAGCAAC 1320
DB 421 SerSerSerGlnAspAlaLeuGlnLysArgLysProSerThrThrLeuAspGlyThrAsn 440
OY 1321 AAGCTATGCAAGTCTCCCTGTTTCAAGAACTATGAGAGAGAGACTCCGGAACAGTCTCT 1380
DB 441 LysLeuGlyGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
OY 1381 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 461 AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
OY 1441 AGCAAGCAATTCATTCGCTCAGAAACAGCAGCAGAGTGGACAGCGCCAGAGTCCCTTTA 1500
DB 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
OY 1501 TCTCCACTGATCAGAAAGTGGAGAGGTTGAGAGCAATTAACAACAGCTTCTTTTCGCG 1560
DB 501 SerProLeuHisAlaGlySerGlySerValGlnAspAsnLysThrSerSerPheLeuPheGly 520
OY 1561 CTTTCCACCAAGCAGCAGACACTCAGCAAGTCTGCTGGCGGCTTTAAAGGCTGGCAC 1620
DB 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlnLysLeuLysGlyThrHis 540
OY 1621 TCGATATCTTGGGCCCCCGAGACCTTACCTTCCCTTCCCTGACAGAGCTGTATTTTGGC 1680
DB 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrPheAla 560
OY 1681 ACAAGTCTTCACTTCTACTGCTCAGCAGCATCTACAGAGAGAGAGAGAGAGAGAGAG 1740
DB 561 ThrGlnSerSerHisPheThrSerAlaSerAlaIleIleArgLysSerAlaSerLysSer 580
OY 1741 GCTTACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
DB 581 AlaThrSerCysSerGlnLeuProThrCysGlyAspGlnValLysSerValArgArg 600
OY 1801 CAGAGAGCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
DB 601 GlnLysProSerAspArgAlaAspSerArgAspSerThrPheHisGlnGlnLysProPheGln 620
OY 1861 AAGCAGTTTAAACGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
DB 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
OY 1921 AGGTCAAGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
DB 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
OY 1981 ATCATTTAGAGGTCTCC 1995
DB 661 IleIleGlnValSer 665
```

RESULT 2
US-09-964-277-2
; Sequence 2, Application US/09964277

```

: Patent No. US20020131770A1
: GENERAL INFORMATION:
: APPLICANT: LucHe, Ralf M.
: APPLICANT: Wei, Bo
: TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
: FILE REFERENCE: 200125_434
: CURRENT APPLICATION NUMBER: US/09/964,277
: CURRENT FILING DATE: 2001-09-25
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 665
: TYPE: prt
: ORGANISM: Homo sapiens
: US-09-964-277-2

```

Alignment Scores:	
Pred. No.:	5.84e-234
Score:	3406.00
Percent Similarity:	99.85%
Best Local Similarity:	99.70%
Query Match:	94.22%
DB:	9
	Gaps: 0

US-10-029-345A-108_COPY_538_2532 (1-1995) X US-09-964-277-2 (1-665)

QY	1	NTGGCCCATGAGATGATTGGAACTCAAAATTTGTTACTGAGAGGTTGGTGGCTGCTGCGAGAA	60
Db	1	MetAlahIsgImetIlegIlythrGlnIleValThhGluArgLeuValAlaLeuLeuGlu	20
QY	61	AGTGAACGGAAAAAGTGCTGCTAAATTGATAGCGCGCCATTGTGTGAAATACATACATCC	120
Db	21	SerGlyThrGluIleValLeuLeuLeuIleAspSerArgProPheValGluIlyrAsnThrSer	40
QY	121	CACATTTTGGAAAGCCATTAAATTCACACTGCTCCAGCTTATGAGACGGAAGGTTGCAACAG	180
Db	41	HsrIleLeuGluAlaIleLeuAsnIleAsnCySerLySLeuMetLyAspArgLeuGlnGln	60
QY	181	GACAAAGTGTAAATTAACAGAGCTCATCCAGACATTCAAGCCAAACATAAAGTTGACATTGAT	240
Db	61	AspLySValIleuIleThrcIuLeuIleGlnIhIAserAluIlyhIbIyValAspIleAsp	80
QY	241	TGCACTCAGAAAGGTTGTAGTTTACGATCAAAGCTCCCAAGATGTGGCTCTCTCTTCA	300
Db	81	CySerGlnLySValIleValIlyrAspGlnSerSerGlnAspValAlaSerLeuSerSer	100
QY	301	GACGTGTTTCTACAGTCACTTCGGGGTAAACAGAAAGAGTTCAACTGTGTACCTG	360
Db	101	AspCySpheLeuThrValLeuLeuGluLySLeuGluIlySerPheAsnSerValhIbLeu	120
QY	361	CTTGCAGGTGGGTTTGCTAGATTCTCTCGTGTGTTTCCGTGGCCTCTGTGAAGGAAATCC	420
Db	121	LeuAlaGlyGlyPheAlaGluPheSerArgCySpheProGlyLeuCySgIuGlyLySer	140
QY	421	ACTCAAGTCCCACTCGCATTTCTCGAGCTTGCTACGTTCACGTGGCCAAACATTGGGCCAAC	480
Db	141	ThrIleuValProThrCySIIeserGlnProCySLeuProValAlaAsnIleGlyProThr	160
QY	481	CGAATTCCTCCCAATCTTTATCTTGAGTCCACAGCGAGATGTCTCTCAACAAGAGCTGATA	540
Db	161	ArgIleLeuProAsnLeuThrIleuGlyCySgInArgAspValLeuAsnLySgIuLeuMet	180
QY	541	CAGCGAATGGGATTTGGTTATGTGTTAAATGCCAGTAAATACCTGTCCAAAGCTGACTTT	600
Db	181	GlnGlnAsnArgLyIleGlyIlyrValLeuAsnAlaSerAsnThrCySProLySProAspPhe	200
QY	601	ATCCCCGAGTCTCATTTCTCGCGGTGGCTGTGAAGACAGCTTTGTGAGAAAATTTTG	660
Db	201	IleProGlnSerHsrPheLeuAsnArgValProValAsnAspSerPheCySgIuLySIIeLeu	220
QY	661	CCGTGTTGAGCAAAATCAGTAGATTTCATTGAGAAAAGCAAAAGCTTCCATGGATGTGT	720
Db	221	ProThrPleuAspLySerValAspPheIleGlyValAlaIlyAlaSerGlnGlyCySVal	240

QY	CTAGTGCACGTTTAGTGGGAGATCTCCGGTCCGGCCACCAATCGGTATCGGCTCATATG	780
Db	LeuValHisCySeuValaGlyIleSerHisSerValaThrIleAlaIleAlaTyrlIleMet	260
QY	AAGAGATGGAACATGTCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAGAAGCTTACT	840
Db	LysArgIgneArbMetSerIleuAerGluAlaTyrlaArgPheValIlySerGluLysArgProThr	280
QY	ATATCTCCAAACTTCATTTTCTGGGGCAACTCTGTGACATATGAGAAAGATTAAACAAC	900
Db	IleSerProAsnPrbAenPheIleGlyIleLeuLeuArpTyrlGlyLysIleLysAsn	300
QY	CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAAGCTGTGCATCGAGAAAGCCAAAT	960
Db	GlnThrIcIylAerGlyProLysSerLysLeuLysLeuLeuHisLeuGluLysProAsn	320
QY	GAACCTGTCCCTGTCTCTCAGAGAGGTGAACAGAAAGCAGACGCCCTCACTCCAGCC	1020
Db	GluProValIProAlaValSerGluGluGlyIleLysSerGluThrProLeuSerProPro	340
QY	TGTGCGCACTGTGTACTCAGAGGGCAGAGCAAAAGGCCGGGCACTCCGCGCAGCGTG	1080
Db	CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArpProValHisProAlaSerVal	360
QY	CCGACGCTGCCACGCTGCAGCCGTCGCTGTAGAGCAGACGCCGCTGTATACAGCGCTC	1140
Db	ProSerAlaIProSerValGlnProSerLeuLeuGlnArpSerProLeuValGlnAlaIleu	380
QY	AGTGGGCTGCACCTGTCTCCGAGACAGCGGTGGAAGACAGCAATAAGCTCAAGCCTTCTT	1200
Db	SerGlyLeuHisIleuSerAlaAspArgLysGluGlnArpSerAlaHisLysLeuLysArgSerPhe	400
QY	TCTGTGATATTCAAATCAGTTTCATATTCATTCAGCCGACATGGCAGACATCCTTACATGCTTC	1260
Db	SerLeuAspIleLysSerValSerTyrlSerAlaSerMetAlaAlaSerLeuHisGlyPhe	420
QY	TCTTCATCAAGAGATGCTTTGAATACTACAAACTTTCACACTACTCTGGATGGAGCCAAAC	1320
Db	SerSerSerGluArpAlaIleuGluTyrlTyrlLysProSerThrIleuAerGlyThrAsn	440
QY	AAGGTATGCGAGTTCTCCCGTTCCAGAACTATCGAGAGAGACTCCGAAACCAAGCTCT	1380
Db	LysLeuCyArgIlePrbSerProValGlnIleLeuSerGluGlnThrProGluThrSerPro	460
QY	GATAGAGAGAGAGCCAGCATCCCCAGAAAGCTGCGACGCGCAGGCTTCCAGACGSCAG	1440
Db	AspLysGluGluIleuAlaSerIleProLysLysLeuGlnThrAlaArgProSerAerSerGln	480
QY	AGCAAGGATTCGACTTGGTTCAGAACCGACAGAGTGGCACCGGCCAGAGTCCCTTTTA	1500
Db	SerLysArgLeuHisSerValaArgThrSerSerSerGlyThrAlaGlnaArgSerLeuLeu	500
QY	TCTCCACTGATCGAAATGGGAGGAGGTGAGAGACAAATTACACACACAGCTTCTTTTGGC	1560
Db	SerProLeuHisIleArgSerGlySerValaGluArpAsnLysHisThrSerPheLeuPheGly	520
QY	CTTTCACACGACCGACGACCTTCCAGAAAGTGTGGCTGGGCTTTAAAGGCTGGCAC	1620
Db	LeuSerThrSerGlnGlnHisIleuThrLysSerAlaGlyLeuGlyLysLysGlyTyrlPheHis	540
QY	TTCGATATCTTGGCCCCCAGACCTTCACCCCTTCCCTGACACGAGCTGGATTTTGGC	1680
Db	SerAspIleLeuAlaProGlnThrSerThrProSerLeuHisSerSerTyrlTyrlPheAla	560
QY	ACAAGATCTCCACACTTCTACTGTGCTGACGCACTTACGAGAGCAGTGCACGATTACTCT	1740
Db	ThrGluSerSerHisIlePheTyrlSerAlaSerAlaIleTyrlGlyLysSerAlaSerTyrlSer	580
QY	GCTACACGCTGCAGCCAGCTGCCACTTGGGGAGACCTAACTGTAATTTGTGCGCAGGCGG	1800
Db	AlaLysIleCySerSerGluLysProThrCySerGlyAspGlnValTyrlSerValaArgArgArg	600


```
QY 1801 CAGAACCAAGTGAAGAGCTGAGCTGGCGGCGAGAGCTGCAATGAGAGACCCCTTTGAA 1860
Db 601 GlnLysProSerAspArgIleAspSerArgArgSerTrpHisGluGluSerProPheGlu 620
QY 1861 AAGCGCTTTAAACGCAAGAGCTGCCAAATGGAAATTTGAGAGAGCAATCATGTACAGAAC 1920
Db 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGluIleMetSerGluAsn 640
QY 1921 AGTCACGCGAAGAGAGCTGGGAGAGTGGGCAAGTCAAGTCTGCTTTTGGGCGAGCATGGAA 1980
Db 641 ArgSerArgGluGluLeuGluLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
QY 1981 ATCATTTAGGCTCTCC 1995
Db 661 ILeIleGluValSer 665
```

RESULT 3
US-10-072-012-680
; Sequence 680, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:

APPLICANT: Tchiernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zethusen, Bryan
APPLICANT: Patlurajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Beba
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Futrak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsbrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 680
; LENGTH: 665

TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-072-012-680
Alignment Scores:
Pred. No.: 5,84e-234
Score: 3406.00
Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 94.22%
DB: 12
Gaps: 0

US-10-029-345a-108_copy_538_2532 (1-1995) x US-10-072-012-680 (1-665)

```
QY 1 ATGGCCCATGAGATGATTTGGAACCAATTGTTACTGAGAGGTGGTGGCTCTGCTGGA 60
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
QY 61 AGTCGAACGGAAGGAGTGGCTTAATTTGATAGCCGAGCATTTGTGGAATACAAATCATCC 120
Db 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyraHisThrSer 40
QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAGGCTTATGAGCGAAGGTTGCAAG 180
Db 41 HisIleLeuGluValIleLeuValIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
QY 181 GACAAAGTGTATTTATTCAGAGCTCATCCAGCATTTGAGCGAAGCAATTAAGTTGACATTGAT 240
Db 61 AspLysValLeuIleIleHisGluLeuIleGlnHisSerAlaTyrHisValAspIleAsp 80
QY 241 TGCAGTCAGAAAGGTGTAGTTTACGATCAAGCTCCCAAGATGTGGCTCTCTCTTCA 300
Db 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTTCACATGTAATCTGCTGTTAACTGGAGAAGAGCTTCAACTCTGTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuGluLysLeuGluLysSerPheAsnSerValHisLeu 120
QY 361 CTTGCAAGTGGGTTTCTGAGTTCTCTGCTGTTGTTTCCCTGGAGCTGTGTAAGAAATTC 420
Db 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlySerGlnLysSer 140
QY 421 ACTTAGTCCCTACCTGATTTCTACGCTTCTTACTGTTGCAAACTTGGGCAAC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCCAGCGAGATGTCTTCAAGAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuLysGlyGlySerGlnArgAspValLeuAsnLysGluLeuMet 180
QY 541 CAGCAGATGAGATGATGTTATGTTTAAATGCGACCTATACCTGTCCAAAGCTGACTTT 600
Db 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAAGTCTCATTTCTGCGTGGCTGCAATGACACTTTTGGAGAAATTTTG 660
Db 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
QY 661 CCGTGGTGGACAAATTCAGTGAATTTCAATTTGAGAAAGCAAAAGCTTCAATGATGTGT 720
Db 221 ProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTATGTCACATGTTAGCTGGAGATCTCCGCTCGCCACCAATGCTTATGCTTACATCAG 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerHisThrIleAlaIleAlaTyrlleMet 260
QY 781 AAGAGATGAGACAGTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACTACT 840
Db 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluLysArgProThr 280
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCATGAGAAAGATTTAAGAC 900
Db 281 IleSerProAsnPheAsnPheLeuGlyLysLeuLeuAspTyrGluLysLysIleLysAsn 300
```

QY 901 CAGACTGAGCATCAGGCGCAAGAGCAACTGAGCTGCTGACCTGAGAGCAACAT 960
 Db 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuH1sLeuGluLysProAsn 320
 QY 961 GAACGTGCTCCCTGCTGCTCAGAGGGTGAGCAGAAAAAGCAGAGCGCCCTCAGTCCACCC 1020
 Db 321 GluProValProAlaValSerGluGlyGlyGlnLysSerLysThrProLeuSerProPro 340
 QY 1021 TGTGGCGACTCTGCTACTCAGAGGCGAGCAGCAAAAGCCCGTGCATCCCGCAGCGTG 1080
 Db 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 QY 1081 CCGAGCGTGGCCGCGCGAGCGCGCTGCTTTAGAGACAGCCCGCTGGTACAGCGGCTC 1140
 Db 361 ProSerValProSerValGlnProSerLeuLysGlnLysPseProLeuValGlnAlaLeu 380
 QY 1141 AGTGGCGTGCACCTGCTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCTTCCTTC 1200
 Db 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerLysLeuLysArgSerPhe 400
 QY 1201 TCTCTGATATCAATATGATTATATATTCAGCAGCATGGCAGCATCTTACATGGCTTC 1260
 Db 401 SerLeuAspIleLysSerValSerLysSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 QY 1261 TCTCTCATCAGAAATGCTTTGGAACTACTACAAACCTTCCACTACTCTGATGGGACCAAC 1320
 Db 421 SerSerSerLysLysPheAlaLeuGluLysLysLysProSerThrThrLeuAspGlyThrAsn 440
 QY 1321 AAGCTATGCGAGTCTTCCTGCTTCCAGGAGACTATCGGAGAGAGATCCCGCAACCAAGCTC 1380
 Db 441 LysLeuGlySerGlnPheSerProValGlnGlnLeuSerGlnThrProGlnThrSerPro 460
 QY 1381 GATAGAGGAGAGAGCCAGCATCCCAAGAGCTGCAGACCGCCAGCGCTTCAGACAGCCAG 1440
 Db 461 AspLysGluGlnLysLeuSerLysLeuThrProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 QY 1441 AGCAGAGGATGCTGATTTGGTTCAGAACCAAGCAGAGCTGGCCCGCCAGAGAGTCCCTTTTA 1500
 Db 481 SerLysArgLysLeuHisSerValArgThrSerSerSerLysThrAlaGlnAspSerLeuLeu 500
 QY 1501 TCTCCACTGATCGAAGTGGAGGAGGCTGAGAGCAATACCAACACAGCTTCCTTTGCGC 1560
 Db 501 SerProLeuHisArgSerGlySerValGlnAspAsnThrHisThrSerPheLeuPheGly 520
 QY 1561 CTTTCCACGAGCAGCAGACCTTCAGAGCTGTGCTGGCGCTTAAAGGCGTGGCAC 1620
 Db 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyThrHis 540
 QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACCAAGAGCTGGTATTTGGC 1680
 Db 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerLysThrPheAla 560
 QY 1681 ACAGAGCTCTACACTTCTACTGCTCAGCCACTTACCGAGGAGGAGGAGGAGGAGTACTCT 1740
 Db 561 ThrGlnSerSerHisPheLysThrSerAlaSerAlaIleLysGlyLysSerAlaSerLysSer 580
 QY 1741 GCCTACAGCTGACAGCCAGCTGCCACTTGGCGAGACCAAGCTTATCTGTGCGAGCGG 1800
 Db 581 AlaLysSerCysSerGlnLeuProThrCysGlyAspGlnValLysSerValArgArgArg 600
 QY 1801 CAGAGGCAAGTACAGAGCTGACTCGCGGCGAGAGCTGGCATGAGAGAGAGAGAGAGAGAG 1860
 Db 601 GlnLysProSerAspArgAlaAspSerArgArgSerLysIleGluGlnLysProPheGln 620
 QY 1861 AAGCAGTTTAAACGACAGAGCTGCCAATGAAATTTGAGAGAGAGCATGTCAGAGAAC 1920
 Db 621 LysGlnLysPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerLysLysMetSerGluAsn 640
 QY 1921 AGGTACGCGGAAAGCTGGGAGAAAGTGGGAGCTGACGTGACTTTTGGGCGAGAGTGA 1980
 Db 641 ArgSerArgGlnGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660

QY 1981 ATCATGAGGTCTCC 1995
 Db 661 IleIleGlnValSer 665
 RESULT 4
 US-10-168-506-14
 ; Sequence 14, Application US/10168506
 ; Publication No. US20040053229A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GREGORY D.
 ; APPLICANT: MARTINEZ, RICARDO
 ; APPLICANT: WHYTE, DAVID
 ; APPLICANT: MANNING, GERARD
 ; APPLICANT: SUDASANAM, SUCHA
 ; APPLICANT: HILL, RON
 ; APPLICANT: FLANAGAN, PETER
 ; TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
 ; FILE REFERENCE: 038602/1351
 ; CURRENT APPLICATION NUMBER: US/10/168,506
 ; PRIORITY FILING DATE: 2002-06-21
 ; PRIOR APPLICATION NUMBER: PCT/US00/34736
 ; NUMBER OF SEQ ID NOS: 76
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 665
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-168-506-14
 Alignment Scores:
 Pred. No.: 5 84e-234 Length: 665
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 94.22% Indels: 0
 Gaps: 0
 US-10-029-345a-108_COPY_538_2532 (1-1995) x US-10-168-506-14 (1-665)
 QY 1 ATGGCCCATGAGATGATGGAATCTCAATGTTTACTGAGAGGTGGCTGCTGCGAA 60
 Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGln 20
 QY 61 AGTGAACGGAAGAAAGTGTGCTGATTTGATGATGCGCGGCACTTTTGGAAATTCATATCC 120
 Db 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluLysThrSer 40
 QY 121 CACATTTGGAAGCCATTATATCACTGCTCCAGAGCTTATGAAAGGAGTTGCAACAG 180
 Db 41 HisIleLeuGlnAlaIleLeuIleAsnLysCysSerLysLeuMetLysArgArgLeuGln 60
 QY 181 GACAAAGTGTATATTCAGAGCTCATCAGACTTCAGCGAAACATAGGTTGACATTGAT 240
 Db 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
 QY 241 TGCAGTCAGAAAGTGTGATTTACATCAAGTCCCAAGATGTGCTCTCTCTTCA 300
 Db 81 CysSerGlnLysValValValLysArgGlnSerSerGlnAspValAlaSerLeuSerSer 100
 QY 301 GACTGTTTCTCAGTCTGATCTGAGGTAAATCGAAGAGGCTTCAACTGTGTCACCTG 360
 Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
 QY 361 CTTCAGAGTGGGTTTGCTGAGTTCTGCTGTTGTTTCCCTGGCCTGTGAAGAGAAATCC 420
 Db 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlnLysSer 140
 QY 421 ACTCTGATCCCTACCTGCAATTTCTGAGCCTTGTTACCTGTTGCCAACTATGGGCCA 480
 Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 QY 481 CGAATCTTCCCAATCTTATCTTGCTGCGCAGAGAGATGTCTCAACAAAGAGTGTATA 540

```
Db 161 ArgIleuProAsnLeuIleuGlyCysGlnIleuAspValLeuAsnIleuMet 180
QY 541 CAGCAGAAAGGATGTTGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCCTGACTTT
Db 181 GlnGlnAsnGlyIleGlyIleValLeuAsnIleAsnThrCysProIleProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCCGTGCTGCTGAGAAATGACAGTTTGTGAGAAAATTTTG
Db 201 IleProGlnSerHisPheLeuArgValProValAsnMetSerPheCysGlnIleLeu 220
QY 661 CCGTGTGAGCAAAATCAGTACATTTCAATGAGAAAGCAAAAGCCTCCATGATGTTGT
Db 221 ProIleLeuAspIleSerValAspPheIleGlnIleGlnIleValIleAsnGlyCysVal 240
QY 721 CTATGCACTGTTTGTGAGGATCTCCGCTCCGCAACATCCGCTATCCGCTATCATG
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
QY 781 AAGAGATGACATGCTTTTATGATGAGCTTACAGATTTGTGAAAGAAAAGACCTACT
Db 261 LysArgMetAspMetSerLeuAspGlyAlaIleArgPheValIleGlnIleValArgProThr 280
QY 841 ATATCTCCAAACTTCATTTTCTGAGGCAATCTCTGACATATGAAAGAAATTAAGAA
Db 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuAsnAspIleGlnIleValIleValAsn 300
QY 901 CAGACTGAGGATCAGAGGCGCAAGAGCAAACTCAAGCTGTGACCTGGAGAAAGCAAA
Db 301 GlnThrGlyAlaSerGlyProIleSerIleuIleuIleuIleuIleuIleuIleuIleu 320
QY 961 GAACCTGCTCCGCTGTCTAGAGGGTGAACAGAAAAGCAGAGCGCCCTCAGTCCAGCC
Db 321 GluProIleProAlaValSerGlnIleGlyGlnIleSerGlnIleProIleuSerProPro 340
QY 1021 TGTGCGCACTGTGTACTCAGAGGAGCAGAGCAAAAGCGCTGATCCGCGCAGCTTG
Db 341 CysIleAspSerAlaThrSerGlnIleAlaGlyGlnArgProValHisProIleAsnVal 360
QY 1081 CCCAGCGTGCAGCGGCGGCGCGCTGTTTGAAGACAGCCCGCTGTGACAGCGCTC
Db 361 ProSerValProSerValGlnProSerIleuLeuGlnIleAspSerProLeuValGlnIleLeu 380
QY 1141 AGTGGCTGACCTGTCCGAGACAGAGCTGAGAAAGCAGAAATTAACCTTAAGCTTCC
Db 381 SerGlyLeuHisIleuSerAlaAspArgLeuGlnIleAspSerAsnIleuIleuIleuIleu 400
QY 1201 TCTGTGATATCAAAATCACTTTCATTTTACGCCAGCATGGAGCATCCCTTACATGCTTC
Db 401 SerLeuAspIleIleuSerValSerIleSerAlaSerMetAlaIleSerLeuHisGlyPhe 420
QY 1261 TCTCATCAGAAAGATGCTTGTGAAATCTAACAACCTTCCACTCTGAGTGGAGCAAC
Db 421 SerSerSerGlnIleAspAlaLeuGlnIleIleIleIleIleIleIleIleIleIleIle 440
QY 1321 AAGCTATGCGACTTCCCTGTTTCAAGAACTATTCAGAGCAGACTCCGAAACCAAGTCT
Db 441 LysLeuCysGlnPheSerProValGlnIleuSerGlnIleuProGlnIleuIleuIleu 460
QY 1381 GATAAGAGAGAAAGCAGAGATCCCAAGAAAGCTGAGACCGCGAGGCTTCAAGACGCA
Db 461 AspIleGlnIleGlnIleSerIleProIleIleuGlnIleThrIleArgProSerAspSerGln 480
QY 1441 AGCAAGCATGTCATTCGGTCAAGAACAGCAGAGTGGCAACCGCCAGAGAGTCCCTTTA
Db 481 SerIleAspArgLeuHisSerValArgThrSerSerSerIleIleThrAlaGlnIleGlnSerLeu 500
QY 1501 TCTCAGCTGATGCAAGTGGAGAGCTGAGAGCAATTAACAACAGCTTCTTTTGGC
Db 501 SerProLeuHisArgSerGlySerValGlnIleAspAsnIleIleThrSerPheLeuPheGly 520
QY 1561 CTTTTCACAGGCGAGGACCTCAGCAAGTCTGTGCTGGCTGAGGCTTAAAGGCTTGAC
1620
```

```
Db 521 LeuSerThrSerGlnGlnHisIleuThrIleuSerAlaGlyLeuIleuIleuIleuIleu 540
QY 1621 TGGATATTTTGCCCCCGAAGACTTACCCCTTCCCTGACAGACAGCTGATTTTGGC
Db 541 SerAspIleLeuAlaProGlnIleThrSerThrProSerIleuThrSerSerIlePheAla 560
QY 1681 ACAGAGTCCCTACACTTCTACTCTGCGCTCAGGCAATCTACAGGAGGAGGAGTCTTCT
Db 561 ThrIleuSerSerHisPheIleuSerIleSerAlaIleIleIleIleIleIleIleIleIle 580
QY 1741 GCTTACACTGACAGCAGCTGCCCACTTTCGAGAGACCAAGTCTATTTCTGTGCGCAGCG
Db 581 AlaIleSerCysSerGlnIleuProThrCysGlyAspIleValIleSerValAlaArgAlaArg 600
QY 1801 CAGAAAGCAAGTGCAGAGCTGACTCGCGCGGAGAGCTGGATGATGAAGAGCCCTTTGAA
Db 601 GlnIleProSerAspAlaAlaAspSerAlaArgSerIleIleGlnIleuSerProPheGln 620
QY 1861 AAGCATTTAAACGACAGACTGCGCAATGGAATTTGAGAGAGCATCATGTCAAGAAC
Db 621 LysGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 640
QY 1921 AGTCAAGGAGAGAGCTGGGAGAAAGTGGAGATGATGATGATTTTGGGCGAGCATGAA
Db 641 ArgSerArgIleGlnIleuGlyIleValIleGlySerGlnIleSerPheSerIleuMetGln 660
QY 1981 ATCATTGAGTCTTC 1995
Db 661 IleIleGlnValSer 665
```

RESULT 5
US-10-343-357-7
Sequence 7, Application US/10343357
Publication No. US20040058341A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom
APPLICANT: BILLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
APPLICANT: YAO, Monique G.; BURROUD, Neil
APPLICANT: WANG, Yumel E.; STEWART, Elizabeth A.
APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
APPLICANT: LEE, Ernestine A.; HAFALIA, April J. A.
APPLICANT: LU, Dying Alina M.; TRIBOULET, Catherine M.
APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
APPLICANT: YUE, Henry; WARREN, Bridget A.
APPLICANT: NGUYEN, Daniel B.; CHAMLA, Nirinder K.
APPLICANT: KEARNEY, Liam
TITLE OF INVENTION: PROTEIN PHOSPHATASES
FILE REFERENCE: PI-0173 PCT
CURRENT APPLICATION NUMBER: US/10/343,357
PRIOR FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: PCT/US01/23716
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,679
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/223,272
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/224,309
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 60/226,728
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 60/229,254
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: US 60/231,366
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CD1

US-10-343-357-7

Alignment Scores:
Pred. No.: 5.84e-234 Length: 665
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 94.22% Indels: 0
DB: 12 Gaps: 0

US-10-029-345a-108_copy_538_2532 (1-1995) x US-10-343-357-7 (1-665)

QY 1 ATGGCCCATGAGTGGATGGACCAATGTTACTGAGAGTGGTGGCTGCTGGAA 60
DB 1 MetAlaHisGluMetCileGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGlu 20
QY 61 AGTGAACGGAAGAGTGTCTTAATTTGATAGCCGGCATTGTGGAAATCAATACC 120
DB 21 SerGlyThrGlnIleValLeuLeuIleAspSerArgProPheValGlnIleThrSer 40
QY 121 CACATTTTGAAGCCATTAAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACAG 180
DB 41 HisIleLeuGlnIleValIleAsnIleAsnCySerLysLeuMetLysArgArgLeuGln 60
QY 181 GACAAAGTGTATTAACAGCTCATCCAGCATTCAGCGAACAATAGTTGACATTGAT 240
DB 61 AspLysValIleuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY 241 TGCAGTCGAAGGTTGATTGTTACGATCAAAAGCTCCCAAGATGTGGCTCTCTTCA 300
DB 81 CysSerGlnIleValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTTCACGTGATCTTGCGGTAACTGGAGAAAGCTTCAACTGTTCACCTG 360
DB 101 AspCysPheLeuThrValIleuLeuGlyLysLeuGlnIleuSerPheAsnSerValHisLeu 120
QY 361 CTGGAGGTTGGTGGTGGTCTGCTGCTGTTGTTCCCTGGCTGGTGAAGGAAATCC 420
DB 121 LeuAlaGlyLysPheAlaGlnPheSerArgCysPheProGlyLeuGlyLysSer 140
QY 421 ACTGATCCCTACCTGACATTTCTCAGCCTTGTCTTACCTGTGCAACATTTGGCCAAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTTCCCAATCTTTATCTTGCGTCCAGCGAGATGCTTCAACAAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValIleuAsnLysGlnLeuMet 180
QY 541 CACGAGAATGGGATGGTGTATGTTAAATGCCAGCTATACCTGTCGAAGCCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCCAGATCTCATTTCTGCGGTGCTGCTGATAGACAGCTTTTGTGAGAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIleLeu 220
QY 661 CCGTGTGGACAATAGTAGATTTCAATGAGAAACAAAGCCTTCAATGATGTGT 720
DB 221 ProThrPheAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTGTGACATGTTAGCTGGGATCTCCGCTCCGCGCAACATTCGCTATCGCTCATCANG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerIleAsnIleAlaIleAlaTyrIleMet 260
QY 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGTGAGAAAGAAACCTACT 840
DB 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlnIlePheProThr 280
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnIleLysIleLysAsn 300
QY 901 CAGACTGAGCATCAGGCGCAAGACAACTCAAGCTGCTGCACTGGAGAAACCAAT 960

DB 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlnIleLysProAsn 320
QY 961 GAACTGTCTCCCTGCTGTCTCAGAGGCTGACAGAAAGCGACAGCCCTCACTCAACC 1020
DB 321 GlnProValProAlaValSerGlnIleGlyGlnIleLysSerGlnIleThrProLeuSerPro 340
QY 1021 TGTGGCAGCTGTCTACCTCAGAGGAGAGGCAAGAGCCGCTGATCCCGCAGCGTG 1080
DB 341 CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal 360
QY 1081 CCCAGCGTCCCAAGCTGAGCGCTGCTGCTTTAGAGACAGCCCGCTGTACAGCGCTTC 1140
DB 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGCTGCACTGTCCGCAACAGCTGGAAAGACAGCAATAGCTCAAGCTTCTCTTC 1200
DB 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 400
QY 1201 TCTGTGATATGAAATCATTTTCAATTAATGAGCAGATGGCAGCATCTTACATGGCTTC 1260
DB 401 SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
QY 1261 TCTCATCAGAAAGATGCTTTGGAATCTACAAACCTTCCACTACTGTGATGGAGCAAC 1320
DB 421 SerSerSerGlnAspAlaLeuGlnIleTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
QY 1321 AAGCTATGSCAAGTCTTCCCTGTTCAAGAACTATCGAGACAGCTCCGAAACAGTCTCT 1380
DB 441 LysLeuCysGlnPheSerProValGlnGlnLeuSerGlnIleThrProGlnThrSerPro 460
QY 1381 GATTAAGAGAAAGCGCATCCCAAGAGCTGACAGCGCAGCGCTTCAGACAGCGAG 1440
DB 461 AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
QY 1441 AGCAAGCATTCATTCGCTCAGAACACAGACAGAGTGGACCGCCAGAGATCCCTTTTA 1500
DB 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
QY 1501 TCTCCATCGATCGAAGTGGAGAGCTGGAGAGACAAATTACACACCACTTCTTTTGGC 1560
DB 501 SerProLeuHisArgSerGlySerValGlnAspAsnTyrHisThrSerPheLeuPheGly 520
QY 1561 CTTTCCAGCAGCAGAGACCTCAGAGAGTGTGCGCTGGGCGCTTAAAGGCTGGCAC 1620
DB 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheHis 540
QY 1621 TCGATATCTTGGCCCCCGACAGCTTACCCCTTCCCTGACACGACGCTGTGATTTTGGC 1680
DB 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrTyrPheAla 560
QY 1681 ACAGAGTCTTCACTTACTTCTGCTTACGCCATTTACGAGAGCGAGTGCAGTTACTCT 1740
DB 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSer 580
QY 1741 GCTTACAGCTGACCGACGCTGCCACTTGGGAGACCAAGTCAATCTGTGGCAGGCGG 1800
DB 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
QY 1801 CAGAAACCAAGTACAGAGCTGACTGCGCGGAGCTGGCATGAGAGAGAGCCCTTTGAA 1860
DB 601 GlnLysProSerAspArgAlaAspSerArgArgSerThrPheGlnIleuSerProPheGln 620
QY 1861 AAGCAGTTTAAACGAGAAAGCTGCCAAATGAAATTTGAGAGAGCATCATGTCAAGAAC 1920
DB 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGTCAACGAGGAGAGCTGGGGAAGTGGGCGAGCGAGCTAGCTTTTGGGCGAGCATGAA 1980
DB 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerLysSerGln 660
QY 1981 ATCATTTAGAGTCTCC 1995

Db 661 llelgluValSer 665

```

RESULT 6
US-10-377-072-26
; Sequence 26, Application US/10377072
; Publication No. US20040009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Miyoung
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; FILE REFERENCE: AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/10/377,072
; PRIOR FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 665
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-377-072-26

```

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5	846-234	665	3406	00	663	1	0
Percent Similarity:	99.85%						
Best Local Similarity:	99.70%						
Query Match:	94.22%						

US-10-029-345a-108_copy_538_2532 (1-1995) x US-10-377-072-26 (1-665)

```

QY 1 ATGGCCATGAGCATTTGGAATCTCAATGAGGTTGGTGGCTGCTGGA
Db 1 MetAlahsIsglUmetllEgLyThrglnllEvaIthrglUkrglEuvAlAlaleuLeugln 20
QY 61 AGTGAACGAAAAAGTGTCTTAATTGATAGCGCGCATTTGTGGAATACATATCATCC 120
Db 21 SerGlyThrglnUysValleuLeuIleAspSerAtrProPheValGlnUyZasnThrsr 40
QY 121 CACATTTTGAAGCATTTATATCACTGCTCCAAAGCTTATAGAGGAAGGTTGCAACAG 180
Db 41 HislleuGlUalIleAenlleAenCysSerlyLeuMetlyAhrAhrglEudlnGln 60

```

```

QY 181 GACAAAGTGTATATACAGAGCTCATCCAGATTCCAGCAAAATAGTTGACATTGAT 240
Db 61 AspLySvalleuIleThrglUleuIleGlnHsSerAlaYshlslySvalAspIleAsp 80
QY 241 TGCAGTCAGAGAGTGTGTGTTTACGATCAAGAGCTCCCAAGATGTGCTCTCTTCA 300
Db 81 CysSerGlnUysValIleValIleValIleAspIleSerSerGlnAspValAlaSerLeuSer 100
QY 301 GACTGTTTCTCACTGTATCTTGGGTAACCTGAGAGAGAGCTTCAACTGTGCACTG 360
Db 101 AspCysPheLeuThrValleuLeuGlyUysSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGGTGGGTTTGGCGAGTCTCCGCTGTTTCCCTGCGCTGTGAAGAAATCC 420
Db 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyUysSerGlnGlySer 140
QY 421 ACTGTATCCCTTACCTTATCTTCAAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 141 ThrleuValProThrCysIleSerGlnProCysleuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCATCTTATCTTGGCTGCGCAGCGAGATGCTTCAACAGAGCTGATA 540
Db 161 ArgIleleuProAsnleuUyruenGlyCysGlnArgAspValleuAsnlyGlnUmet 180
QY 541 CAGCAGAAATGGAGTTGTTATGTTAAATGCGACCTATACCTGCTCCAAAGCTGACTT 600
Db 181 GlnGlnAsnIlyIleGlyValleuAsnAlaSerAenthrcySprolySproAspPhe 200
QY 601 ATCCCGAGTCTCATTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 201 IleProGlnSerHisPheleuArgValProValAsnSpsSerPheCysGlnUysIleLeu 220
QY 661 CCGTGTGTGCAAAATCGATGATTTCAATGAGAAGCAAAAGCTCCAAATGATGTT 720
Db 221 ProThrleuAspUysSerValAspPheIleGlnUysAlaYshlslySval 240
QY 721 CTATGCACTGTTTGTGCTGAGTCCCGCTCGGCAACATCGCTATCGCTTACATCATG 780
Db 241 LeuValHisCysleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 260
QY 781 AGAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTT 840
Db 261 LysArgMetAspMetSerleuAspGlnAlaTyArgPheValIlySlnUysArgProThr 280
QY 841 ATATCTCCAACTCATTTTCTGGGCAACTCTCTGACTATGAGAAGATTAAGAAC 900
Db 281 IleSerProAsnPheAsnPheleuGlyGlnleuLeuAspUyrglnUysIlySln 300
QY 901 CAGACTGAGCATCAGGCGCAAGCAAACTCAAGCTTGTGCTGCTGCTGCTGCTGCTGCT 960
Db 301 GlnThrIlyAlaSerGlyProUysSerlySleuUySleuLeuHisleuGlnUysProAsn 320
QY 961 GAACCTGCTCCCTCTGCTTCAAGAGGTGAGAGAAAGCAAGCGCCCTGATCAAGCC 1020
Db 321 GluProValProAlaIleSerGlnUyGlnUysSerGlnUytrProLeuSerProPro 340
QY 1021 TGTGCCGACTCTGCTTACTTCAAGGCGAGCAAGCAAAAGCGCTGATCCGCGCAGGCTG 1080
Db 341 CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal 360
QY 1081 CCCAGCGTCCAGAGGTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 361 ProSerValProSerValGlnProSerleuUyGlnUysSerProleuValGlnAlaLeu 380
QY 1141 AGTGGGCTGACCTTCTCGCAGACAGAGGTGAGACAGATAGCTCAAGCTTCTTCC 1200
Db 381 SerGlyleuHisSleuSerAlaAspArgleuGlnUysSerUySleuUyArgSerPhe 400
QY 1201 TCTCTGATATCAATAGTATTATTCAGCGAGATGCGACATCTTATACATGCTTC 1260
Db 401 SerleuAspIleUysSerValSerlySerAlaSerMetAlaIleSerleuHisGlyPhe 420
QY 1261 TCCATCATAGAAAGTCTTGAATCTCAAACTTCACTACTGTGATGGACCAAC 1320

```

```

Db      421 SerSerSerGlnuAphAlaLeuGluTyrTyrTyrLeuProSerThrThrLeuAseGlyThrAse 440
QY      1321 AACGTAATGCCAGTTCCTCCCTGTTACAGAACTATCGAGAGCAAGCTCCGAAACCAAGTCTT 1380
Db      441 LysLeuCyseGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
QY      1381 GATTAAGAGGAACCCAGCATCCCGCAAGAAAGCTGACCGCGCGCTTTCAGACAGCG 1440
Db      461 AspLysGlnGlnuAseThrLeuProLysLysLeuGlnThrAlaArgProSerAsePheSerGln 480
QY      1441 AGCAAGGATTTGATTCGATTCGATCAGAACAGCAGACAGTGCACCGCGCGAGAGTCCCTTTTA 1500
Db      481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnAseSerLeuLeu 500
QY      1501 TCTCCATCGCATCGAAGTGGAGCGTGGAGACAAATTACCAACCAAGCTTCCTTTTCGAC 1560
Db      501 SerProLeuHisArgSerGlySerValGlnAsePheSerThrSerPheLeuPheGly 520
QY      1561 CTTTCCACCGCAGCAGCAGCTTCACAGAAAGTCTGCGCTGGCGCTTAAGGGCTGGCAC 1620
Db      521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysGlyTyrPheHis 540
QY      1621 TCGGATATCTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGAGCTGATATTTGCC 1680
Db      541 SerAsePheLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrPyrPheAla 560
QY      1681 ACAGAGTCTCTACACTTCTGCTGCTGACCACTTACGAGAGAGAGAGTGCAGTACTCT 1740
Db      561 ThrGlnSerSerHisPheThrSerAlaSerAlaIleTyrGlySerAlaSerTyrSer 580
QY      1741 GCCTACAGCTGCGACCGCGCTGCCACTTGGGAGACCAAGTCTATCTGCGCGCAGCGG 1800
Db      581 AlaTyrSerCyseSerGlnLeuProThrCyseGlyAsePheAlaIleTyrSerValAlaArgAse 600
QY      1801 CAGAAAGCAAGTACAGAGCTGACTCGCGGCGGAGCTGCAGATGAAGAGACCCCTTTGAA 1860
Db      601 GlnLysProSerAsePheArgAlaAsePheArgAseThrPheGlnGlnuSerProPheGln 620
QY      1861 AAGCAGTTTAAACGACAGACCTGCCAATGAATTGGAGAGAGCATGATGACAGAAC 1920
Db      621 LysGlnPheLysArgAseCyseGlnMetGlnPheGlyLysSerIleMetSerGlnuAse 640
QY      1921 AGGTACAGGGAAGAGCTGGGGAAGAGTGCAGTCAAGTCTTTTTCGGGCACAGTGGAA 1980
Db      641 ArgSerArgGlnGlnuLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY      1981 ATCATGAGGTCTCC 1995
Db      661 IleIleGlnuValSer 665

RESULT 7
US-10-257-026-2
; Sequence 2, Application US/10257026
; Publication No. US20040086859A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10KDM5
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-257-026-2

Alignment Scores:
Pred. No.: 5,84e-234 Length: 665
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1

```

```

Best Local Similarity: 99.70% Mismatches: 1
Query Match: 94.22% Indels: 0
DB: 16 Gaps: 0

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-10-257-026-2 (1-665)

QY      1 ATGGCCCATGAGATGATGGAATCAAAATTTTACTGAGAGGTGGTGGCTGTGCTGAA 60
Db      1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
QY      61 AGTGAACGGAAGAAAGTGTGCTTAATGATGACGGGCAATTTGTGAATACAAATACATCC 120
Db      21 SerGlyThrGlnLysValLeuLeuIleAsePheArgProPheValGlnTyrAseThrSer 40
QY      121 CACATTTTGAAGCCATTAAATATCACTGCTCCAGCTTATGAAGGGAAGTTGCAACG 180
Db      41 HisIleLeuGlnuAlaIleAseThrLeuAseCyseLysLeuMetLysArgArgLeuGlnGln 60
QY      181 GACAAAGTGTAAATTAACAGAGCTCATCCAGATTCAGCGAAACATTAAGTTGACATTGAT 240
Db      61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAsePheAse 80
QY      241 TGCAGTCAGAAAGTTGATGATTACGATCAAAGCTCCCAAGATGTTGCTCTCTCTCA 300
Db      81 CyseSerGlnuLysValValValTyrAsePheGlnSerSerGlnAsePheValAseLeuSer 100
QY      301 GACTGTTTCTCACTGATCTTGTGGGTAACTGGAAGAAAGCTTCAACTCTGTCACTTG 360
Db      101 AsePhePheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAseSerValHisLeu 120
QY      361 CTTGCAAGTGGGTTTCTGAGTCTCTGTTGTTTCCCTGCGCTCTGTGAGGAAGAAATCC 420
Db      121 LeuAlaGlyGlyPheAlaGlnPheAseArgCysePheProGlyLeuCyseGlnuLysSer 140
QY      421 ACTTAGTCCCTACCTGCAATTTCTCAGCTTGCTGCTTACTGTTGCCAATTTGGCCCAAC 480
Db      141 ThrLeuValProThrCyseIleSerGlnProCyseLeuProValAlaAseIleGlyProThr 160
QY      481 CGAATTTCTCCCAATCTTTATCTTGGCTGCGACGAGATGTCCTCAACAAAGAGCTGATA 540
Db      161 ArgIleLeuProAseLeuLysLeuGlyCyseGlnArgAsePheAlaLeuAseLysGlnuMet 180
QY      541 CAGCAGAAATGGGATTTGATGTTAAATGCCAGCATATACCTGTCCAAAGCTGACTTT 600
Db      181 GlnGlnAseGlyIleGlyTyrValLeuAseAlaAseThrCyseProLysProAsePhe 200
QY      601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG 660
Db      201 IleProGlnSerHisPheLeuArgValProValAseAsePheCyseGlnuLysIleLeu 220
QY      661 CCGTGGTTGACAAATCAAGTAGATTTCATTGAGAAAGCAAAAGCTCCCAATGAGTGT 720
Db      221 ProThrLeuAsePheSerValAsePheIleGlnuLysAlaLysAseAseGlnuCyseVal 240
QY      721 CTATGTCAGCTTTTACTGGGATCTCCGCTCGGACCAATGCTATGCTTGGCTATCATCATG 780
Db      241 LeuValHisCyseLysuAseGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY      781 AAGAGATGACATGCTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCACTACT 840
Db      261 LysArgMetAsePheSerLeuAsePheGlnuLysArgPheValLysGlnuLysArgProThr 280
QY      841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGACTATGAGAAAGAAAGTTAAGAAC 900
Db      281 IleSerProAsePheAsePheLeuGlnuLysLeuLeuAsePheTyrGlyLysArgIleLysAse 300
QY      901 CAGACTGAGCATCAGAGGCAAGCAAGCAAACTCAAGCTGTGCACTGAGGAAGCAAAAT 960
Db      301 GlnThrGlyAseSerGlyProLysSerLysLeuLysLysLeuLysLysLeuLysProAse 320
QY      961 GAACCTGTCTCTGCTGTCTCAGAGGGTGAACGAAAGGGAAGCGCCCTCACTGACACCC 1020
Db      321 GlnProValProAlaValSerGlnuGlyGlnuLysSerGlnuThrProLeuSerProPro 340

```


QY 1021 TGTCCGACCTCTGCTA CTTCAAGGACGACGACAAAGCCCGCTGATCCCGACGCTG 1080
DB 341 CysAlaAspSerAlaHisSerGlnAlaIleGlnAlaGlnProValHisProAlaSerVal 360
QY 1081 CCCAGGTCCTCCAGCGCTGACCGCTGCTGTTAGAGACGCGCGCTGTTACAGCGCTC 1140
DB 361 ProSerValProSerValGlnProSerIleuGlnuAspSerProIleuValGlnAlaIleu 380
QY 1141 AGTGGGCTGCACCTCTCCGACAGAGCTGAGACGCAATTAAGCTCAAGCTTCCTTC 1200
DB 381 SerGlnIleuHisIleuSerAlaAspArgIleuGlnuAspSerAsnIleuLysArgSerPhe 400
QY 1201 TCTCGAATCAATCAATCAATTCATATTACAGCAGACGACGACCTTAATAGCTTC 1260
DB 401 SerIleuAspIleuLysSerValSerIleuSerAlaSerMetAlaIleuIleuIleuPhe 420
QY 1261 TCTCATCAGAAAGATGCTTTGGAATACAAACCTTCACTCACTGATGAGACCAAC 1320
DB 421 SerSerSerGlnuAspAlaIleuGlnuIleuIleuIleuIleuIleuIleuIleuIleuIleu 440
QY 1321 AAGCTATGCCAGTTCTCCCTGTTTTCAGAACTATGAGACGACCTCCGAAACCACTCT 1380
DB 441 LysIleuGlnIleuHisSerProValGlnIleuSerGlnuIleuIleuIleuIleuIleuIleu 460
QY 1381 GATTAAGGAGGAAAGCCAGCATCCCAAGAGCTGACAGCCGACGCTTCAAGACAGCAG 1440
DB 461 AspLysGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 480
QY 1441 AGCAAGGATTCGATTCGCTGACAAACAGACAGACGACGACGACGACGACGACGACG 1500
DB 481 SerLysArgIleuHisSerValAlaArgHisSerSerGlnIleuIleuIleuIleuIleuIleu 1500
QY 1501 TCTCACTGATCGAATGAGTGGAGCGTGGAGCAATTAACACACAGCTTCCTTTTCGCG 1560
DB 501 SerProIleuHisArgSerGlnSerValGlnuAspAsnIleuIleuIleuIleuIleuIleu 520
QY 1561 CTTTCACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1620
DB 521 LeuSerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 540
QY 1621 TCGAATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACGACGACGACGACGACGAC 1680
DB 541 SerAspIleuIleuAlaProGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1680
QY 1681 ACAGAGCTCACTCACTCTCTGCTTCAAGCATTAAGGACGACGACGACGACGACGACG 1740
DB 561 ThrGlnSerSerHisPheIleuSerAlaSerAlaIleuIleuIleuIleuIleuIleuIleu 1740
QY 1741 GCTTCAAGCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1800
DB 581 AlaIleuSerCysSerGlnIleuProIleuIleuIleuIleuIleuIleuIleuIleuIleu 1800
QY 1801 CAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
DB 601 GlnuIleuProSerAspArgAlaAspSerArgArgSerIleuIleuIleuIleuIleuIleu 1860
QY 1861 AAGCAATTTAAACGAGAGCTGCAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
DB 621 LysGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1920
QY 1921 AGGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
DB 641 ArgSerArgGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 1980
QY 1981 ATCATTTAGGCTCTCC 1995
DB 661 IleIleGlnuValSer 665

RESULT 8
US-10-648-593-240
; Sequence 240, Application US/10648593
; Publication No. US20040106132A1

GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 240
; LENGTH: 665
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-648-593-240

Alignment Scores:
Pred. No.: 5,84e-234 Length: 665
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 94.22% Indels: 0
Gaps: 0

US-10-029-345a-108_copy_538_2532 (1-1995) x US-10-648-593-240 (1-665)

QY 1 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
DB 1 MetAlaHisGlnMetIleGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 20
QY 61 AGTGAACGCAAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
DB 21 SerGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 40
QY 121 CACATTTTGAAGGATTAATATATATATATATATATATATATATATATATATATATATAT 180
DB 41 HisIleuGlnuAlaIleuAsnIleuAsnIleuAsnIleuAsnIleuAsnIleuAsnIleu 60
QY 181 GACAAAGTGAATTAACAGAGCTCATTCAGACGATTCAGGAAACATTAAGGATGACATGAT 240
DB 61 AspLysValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 60
QY 241 TGCAGTCAAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 81 CysSerGlnuValValValValValValValValValValValValValValValValVal 300
QY 301 GACTGTTTCTCATCTGATCTGAGGTAACGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 101 AspCysPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 360
QY 361 CTTGACAGTGGGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
DB 121 LeuAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 420
QY 421 ACTCTAGTCTCTCACTGATTTCTCAGCTTCTTACCTGTTCAACATTTGGGCAACC 480
DB 141 ThrIleuValProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 480
QY 481 CGAATTTTCCCAATCTTATCTTGGCTGACGAGAGATGCTTCAACAGAGGATGATGAT 540
DB 161 ArgIleuIleuProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 540
QY 541 CAGCAGAAATGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB 181 GlnIleuAsnGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 600
QY 601 ATCCCGAGTCAATTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 201 IleProGlnuSerHisPheIleuArgValProValAsnAspSerPheCysGlnuIleuIleu 660
QY 661 CCGTGGTTGAGCAAACTACGATGATTTCAATTTGAGAAAGCAAAAGCTCCAAATGATGATG 720

[illegible][illegible]

```
QY 361 CTTCAGAGTGGGTTGCTGAGTCTCTGTTGTTTCCCTGAGCTCTGTGAGAGAAATCC 420
Db 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlySer 140
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGGCTTGCTTACCTGTTGCAACATTTGGGCCAAC 480
Db 141 ThrLeuValProThrCysGlnSerGlnProCysLeuProValAlaAlaGlnGlyProThr 160
QY 481 CGAATTTCTTCCCAATCTTATCTTGGCTGCCAGGAGATGCTCTCAACAGAGAGCTGATA 540
Db 161 ArgGlnLeuProAlaLeuPheLeuGlyCysGlnArgPheValLeuGlnGlyLeuMet 180
QY 541 CAGCAGATGGGATTTGTTATGTTAAATGCGAGCTATACCTGTTGCAAGAGCTTCTT 600
Db 181 GlnGlnGlnGlyGlyGlyValLeuAlaAlaSerGlnThrCysProArgPhe 200
QY 601 ATCCCGAGTCTCATTTCTGCGTGGCTGTAATGACAGCTTTTGTGAGAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAlaAspSerPheCysGlnGlyLeu 220
QY 661 CCGTGGTTGCAAAATCAGTAGATTTTCATTGAGAAAGCAAGCTTTTGCATGATGTT 720
Db 221 ProThrLeuAspLysSerValAspPheIleGlnGlySalAlaLysAsnGlyCysVal 240
QY 721 CTAGTCACTGTTTACCTGAGATCTCCGCTCCGCACTGCTATGCTATGCTATCATG 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
QY 781 AAGAGATGAGACATGCTTTAGATGAGCTTACAGATTTGTGAGAAAGAAAGAAAGCTT 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaIleArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCCAAATTTCTGAGGCAACTCTGCACTATGAGAAAGATTTAGAAC 900
Db 281 IleSerProAlaPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysIleLysAsn 300
QY 901 CAGATGAGAGATCAGGCGCAAGAGCAAACTCAGGCTGTCAGCTGAGAGAGCGCAAT 960
Db 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisGlnLysProAsn 320
QY 961 GAACCTGCTCCCTGCTCTCAGAGGGTGAACAAGAGAGAGAGCGCTCATGTCACCC 1020
Db 321 GlnProValProAlaValSerGlnGlyGlyGlnLysSerGlnThrProLeuSerProPro 340
QY 1021 TGTCCGCACTGCTACCTCAGAGGAGAGAGCAAGAGCGCTGCATCCGCGCAGGCTG 1080
Db 341 CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal 360
QY 1081 CCGAGCGTGGCCAGCGTGAAGCGGTGTTAGAGAGAGCGCGCTGGTACAGCGCTC 1140
Db 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGCTGCACTGTCGCCCAACAAGGCTGGAAGAAGCAATAAGCTCAGAGCTTCTTC 1200
Db 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAlaLysSerPhe 400
QY 1201 TCTCTGATATCAAAATCAGTTTCATATTCAGCAGAGAGAGATCTTCAATGCTTTC 1260
Db 401 SerLeuAspIleLysSerValSerLysSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
QY 1261 TCTCATCAAGAGAGCTTTGGATATCTCAAACTTTCATCTCTGATGGAGCAAC 1320
Db 421 SerSerSerGlnAspAlaLeuGlnLysTyrLysProSerThrThrLeuAspGlyThrAsn 440
QY 1321 AAGCTATGCCAGTTCTCCCTGTTGAGAGACTATGAGAGAGAGCTCCGAAACAGTCTT 1380
Db 441 LysLeuCysGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
QY 1381 GATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 461 AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480

QY 1441 AGCAGAGATTCATTCGCTGAGAAACAGACAGCTGGACCGGCCAGAGTCCCTTTA 1500
Db 481 SerLysArgLeuHisSerValLysArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
QY 1501 TCTCCATCTGCAATCGAGTGGAGACGCTGAGAGACAATTAACAGACAGCTTCCCTG 1560
Db 501 SerProLeuHisArgSerGlySerValGlnAspAsnThrHisSerPheLeuPheGly 520
QY 1561 CTTTCACAGCAGCAGACACCTCTCAGAGAGCTGCTGCTGCGCTTAAAGGCTGGCAC 1620
Db 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTrpHis 540
QY 1621 TCGGATATCTGCGCCCGCCAGACCTTACCCCTTCCCTGACAGCAGCTGATTTTGC 1680
Db 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpPheAla 560
QY 1681 ACAAGTCTTCACACTTCTACTGCTGCTGAGGCAATTAACGAGAGAGAGCTGCAATTCT 1740
Db 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
QY 1741 GCTTACAGCTGCAAGCCAGCTGCCCACTTGCAGAGACCAAGTCTATTTGTGCGAGCGG 1800
Db 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
QY 1801 CAGAGCGCAATGACAGAGCTGACTCGCGGAGAGCTGGCATGAGAGAGCCCTTTGAA 1860
Db 601 GlnLysProSerAspArgAlaAspSerArgAspSerTrpHisGlnGlnSerProPheGln 620
QY 1861 AACGATTTTAAACGAGAGAGCTGCCAAATGGAATTTGAGAGAGACATCATGAGAGAC 1920
Db 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGGTCAAGGAGAGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
Db 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCATGAGAGCTTCC 1995
Db 661 IleIleGlnValSer 665

RESULT 10
US-10-072-012-679
/ Sequence 679, Application US/10072012
/ Publication No. US2004003493A1
/ GENERAL INFORMATION:
/ APPLICANT: Tchernov, Vellizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Ratturajan, Meera
/ APPLICANT: Shmukets, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangoli, Beba
/ APPLICANT: Padigar, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimiro Y.
/ APPLICANT: Coleman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grose, William M.
/ APPLICANT: Alsobrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Kieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072, 012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
```



```
Qy 1621 TCGATATCTTGGCCCCGAGACTCTACCCCTTCCCTGACAGACCTGGTATTTGGC 1680
Db 566 SerAspIleuAlaLeuAlaProGlnThrThrProSerLeuThrSerSerTrpYrPheAla 585
Qy 1681 ACAGACTCTTCACTTCTTACTCTTGCCTTACAGCCTTACAGAGAGAGAGAGAGAGAGAG 1740
Db 586 ThrGluSerSerHisPheTrpSerAlaSerAlaIleTrpGlyIleSerAlaSerTrpSer 1740
Qy 1741 GCTTACAGCTGAGCAGCAGCTGCCACTTGGCGAGAGACCAAGCTTATTTCTGTGCGAGCGG 1800
Db 606 AlaTrpSerCysSerGlnLeuProThrCysGlyAspIleValTrpSerValArgArgArg 625
Qy 1801 CAGAGCCAGTGAACAGAGCTGACTCGCGCGAGAGCTGCGATGATGAGAGAGAGAGAGAG 1860
Db 626 GlnTrpSerProSerAspArgAlaAspSerArgArgSerTrpHisGlnGlnSerTrpHisGln 1860
Qy 1861 AAGCATTTAAACCCAGAGAGCTGCAATGGAATTTGGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 646 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 665
Qy 1921 AGGTACAGGAGAGAGAGCTGGGAGAAAGTGGCAGTCACTGCTTTCGGGAGAGAGAGAG 1980
Db 666 ArgSerArgGlnGlnLeuGlnLysValGlySerGlnSerSerPheSerGlySerMetGln 685
Qy 1981 ATCATTTAGGCTTCC 1995
Db 686 IleIleGlnValSer 690
```

RESULT 11

US-10-072-012-703

Sequence 703, Application US/10072012

Publication No. US2004003493A1

GENERAL INFORMATION:

APPLICANT: Tcherev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernusen, Bryan
APPLICANT: Patcurajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsbrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767

;; PRIOR FILING DATE: 2001-02-05
;; PRIOR APPLICATION NUMBER: 60/267,057
;; PRIOR FILING DATE: 2001-02-07
;; PRIOR APPLICATION NUMBER: 60/266,975
;; PRIOR FILING DATE: 2001-02-07
;; PRIOR APPLICATION NUMBER: 60/267,459
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1391
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 703
;; LENGTH: 690
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-072-012-703

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5,896-234	3406.00	690	663	1	1	0	0
Percent Similarity:	99.85%						
Best Local Similarity:	99.70%						
Query Match:	94.22%						
DB:	12						

US-10-029-345A-108_copy_538_2532 (1-1995) x US-10-072-012-703 (1-690)

```
Qy 1 ATGGCCCATGATGATGATGGAATGGAATGTTTACTGAGAGTTGGTGGCTGCTGGA 60
Db 26 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuGln 45
Qy 61 AGTGAACCGGAAAGAGTGGTGAATGATGATGATGATGATGATGATGATGATGATGATG 120
Db 46 SerGlyThrGlnValValLeuLeuIleAspSerArgProPheValGlnValAsnThrSer 65
Qy 121 CACATTTGGAAGCATTAATTAATCACTGCTCCCAAGTTTGAAGCGAAGTTGCAACAG 180
Db 66 HisIleLeuGlnAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 85
Qy 181 GACAAAGCTTAATTAATTAATCAAGCTTCATCAGCATTCAGCGAAACATTAAGTTGAT 240
Db 86 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 105
Qy 241 TGCAGTCAGAGGTGTGATGTTTACGATCAAGTCCCAAGATGTTGCTTCTCTTCA 300
Db 106 CysSerGlnLysValValIleValIleValIleValIleValIleValIleValIleVal 125
Qy 301 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAGAGAGTCACTCACTGTTCACTG 360
Db 126 AspCysPheLeuThrValLeuLeuGlnLysLysGlnLysLysSerPheAsnSerValHisLeu 145
Qy 361 CTTCAGAGTGGTGTCTGAGTCTCTGTTGTTTCCCTGGCCTCTGTGAAGAAATCC 420
Db 146 LeuAlaGlyIlePheAlaGlnPheSerArgCysPheProGlyLeuCysGlnGlyLysSer 165
Qy 421 ACTCTAGTCCCTAATCTGATTTCTGAGCTTCTTACCTGTTGCAACATTTGGGCAAC 480
Db 166 ThrIleValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 185
Qy 481 CGAATCTTCCCAATCTTATCTTGGGCGCAGAGAGATGCTTCAACAAAGAGCTGATA 540
Db 186 ArgIleLeuProAsnLeuIleLysIleCysGlnIleAspValLeuAsnLysGlnLeuMet 205
Qy 541 CAGGAGATGGATGGATTGTTGTTAAATGCAAGCTTACTGCTGCAAGAGCTGATA 600
Db 206 GlnGlnAsnGlnIleGlyIleValLeuAsnAlaSerAsnThrCysProLysProAspPhe 225
Qy 601 ATCCCGAGTCTCATTTCTGCGTGTGCGTGAATGCAAGCTTTTGTGAAGAAATTTTG 660
Db 226 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 245
Qy 661 CCGTGGTTGGAACAAATGATGATTTCAATGAGAAAGCAAAAGCTTCAATGATGTT 720
Db 246 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaIleValAsnIleCysVal 265
```

QY 721 CTAGTGACGTTAGTGGGATCTCCGCTCCGCCACATCGGTATCGCTACATCANG 780
DB 266 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 285
QY 781 AAGAGATGACATGTTTGTAGTGAAGCTTACAGATTGTGAAAGAAAAGAACTACT 840
DB 286 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIlySerGluValArgProThr 305
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAATCTCTGACATATGAAAGAAAGATTAAAGAAC 900
DB 306 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluIlyValIleLysAsn 325
QY 901 CAGACTGAGACATCAGGCGCCAAAGAGCAATCAAGCTGTGACCTGAGCGAAGCCAAAT 960
DB 326 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGluLysProAsn 345
QY 961 GAACCTGTCCCTGTCTGACAGAGGTGACAGAAAAGCAGACGCCCCCTCAGTCCACCC 1020
DB 346 GluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProPro 365
QY 1021 TGTGCGCACTCTGTCTACCTCAGAGGCGACGACAAAGGCCCTGATCCGCCAGCGTG 1080
DB 366 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 385
QY 1081 CCGAGGCTGCCCGCCGCGACGCGCTGCTTTAGAGACACGCCCTGTGTACAGCGGCTC 1140
DB 386 ProSerAlaProSerValGlnProSerLeuGlnLysPheProLeuValGlnAlaLeu 405
QY 1141 AGTGGGTGACCTGTCCGACAGAGGCTGGAAGACAGCAATAGCTCAAGCTTCCTTC 1200
DB 406 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 425
QY 1201 TCTCTGATATCAATAGTTTCATATTCAGCAGACATGGCAGCATCTTACATGGCTTC 1260
DB 426 SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 445
QY 1261 TCTCTATCAGAAATGCTTTGGAATATCTACAAACCTTCCACTCTCTGTGATGGGACCAAC 1320
DB 446 SerSerSerGlnAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAsn 465
QY 1321 AAGCTATGCGAGTTCTCCCTGTTCAGGAATCTACGAGAGAGACTCCCGAAACAGACTCT 1380
DB 466 LysLeuCysGlnHisSerProValGlnGlnLeuSerGlnHisProGlnThrSerPro 485
QY 1381 GATTAAGAGAGAACCCAGCATCCCAAGAGCTCAGACCGCCAGGCTTTCAGACAGCAG 1440
DB 486 AspLysGluGlnIleAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 505
QY 1441 AGCAAGCGATTCATTCGGTCAAGACCAAGCAGAGTGGACCGGCCAGAGAGTCCCTTTTA 1500
DB 506 SerLysArgLeuHisSerValArgGlnHisSerSerGlyThrAlaGlnHisSerLeuLeu 525
QY 1501 TCTCCACTGATCGAAGTGGAGCGTGGAGACAAATTACACACAGACTCTCTTTCCGCG 1560
DB 526 SerProLeuHisArgSerGlySerValGlnAspAsnTyrHisThrSerPheLeuPheGly 545
QY 1561 CTTTCCACGACGACGACACTCAGAGAGTGTCTGGCTGGGCTTAAAGGCTGGAC 1620
DB 546 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysValTyrPheHis 565
QY 1621 TCGGATATCTGGGCCCCCGACACTCTACCCCTTCCCTGACGACAGCTGGATTTTGGC 1680
DB 566 SerAspIleLeuAlaProGlnHisThrProSerLeuThrSerSerTyrTyrPheAla 585
QY 1681 ACAGAGCTCTACACTTCTACTGTGCTCAGCCATCTACGAGAGGAGCTGACATTAATCT 1740
DB 586 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLeuSerAlaSerTyrSer 605
QY 1741 GCTTACAGCTGACGACGCTGCCACTTGGCGAGACCAAGCTTAATCTGTGCGCAGCGG 1800
DB 606 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValAlaArgArg 625

QY 1801 CAGAACCAAGTACAGACAGCTGACTCGCGCGAGCTGCGCATAGAAAGACCCCTTTGAA 1860
DB 626 GlnLysProSerAspArgAlaAspSerAlaArgSerTyrPheGlnLeuSerProPheGln 645
QY 1861 AAGCACTTTAAACGCAAGAGCTGCGCAATGAAATTTGAGAGAGCATGTCAGAGAAC 1920
DB 646 LysGlnPheLysArgTyrSerCysGlnMetGluPheGlyGlnSerIleMetSerGluAsn 665
QY 1921 AGGTCAAGGAGAAAGCTGGGGAAGAGTGGGAGTCAATCTAGCTTTCCGGCAGCATGAA 1980
DB 666 ArgSerArgGlnGlnLeuGlyValGlySerGlnSerSerPheSerGlySerMetGlu 685
QY 1981 ATCATGAGGTCCTCC 1995
DB 686 IleIleGluValSer 690

RESULT 12
US-10-425-114-54204
; Sequence 54204, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5331)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54204
; LENGTH: 690
; TYPE: PRN
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-028-H6_F11.dep
; US-10-425-114-54204

Alignment Scores:
Pred. No.: 5,89e-234 Length: 690
Score: 3406.00 Matches: 663
Percent Similarity: 99.858 Conservative: 1
Best Local Similarity: 99.704 Mismatches: 1
Query Match: 94.224 Indels: 0
DB: 12 Gaps: 0

US-10-029-345a-108_copy_538_2532 (1-1995) x US-10-425-114-54204 (1-690)
QY 1 ATGGCCCATGAGATGATGGAATCTCAATTTGTTACTGAGAGTTGGTGTCTGTGGA 60
DB 26 MetAlaHisGluMetIleGlyThrGlnIleValIThrGluArgLeuValAlaLeuLeuGln 45
QY 61 AGTGAACCGGAAAGAGTGTCTGTAATGATAGCCGGCCATTTGTGGAATTAATATATCC 120
DB 46 SerGlyThrGlnLysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 65
QY 121 CACATTTGGAAGCCCTTATATATCAATCTGCTCAAGACTTATGAAAGGAGGTTGCAACAG 180
DB 66 HisIleLeuGlnIleAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 85
QY 181 GACAAAGTGTATTAATACAGAGCTCATCGACATTCAGCAAAACATAGGTTGACATTGAT 240
DB 86 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisIleValAspIleAsp 105
QY 241 TGCAGTCAGAAAGTTGATTAATTAAGATCAAAAGCTCCCAAGATGGTCCCTCTCTTCA 300
DB 106 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspAlaAlaSerLeuSerSer 125
QY 301 GACTGTTTCTCACTGTAATCTTGGGTAAATCGAAGAGCTTCAACTGTTCACCTG 360


```
Db      126 AspCysPheLeuThrValIleuLeuGlyIysLeuGluIysSerPheAsnSerValIhIleu 145
QY      361 CTTCGAGGTGGGTTTGCTGAGATTCTCTGTTGTTTCCCTGCGCTCTGTGAGAAATCC 420
Db      146 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyIysSer 165
QY      421 ACTTCATGTCCTACCTGATTCTCAGCCCTGCTTACTCTGTGCAACATTTGGCCAC 480
Db      166 ThrIleuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 185
QY      481 CGAATTTCTCCCAATCTTTATCTTGCTGCTCCAGAGAGATGCTCTCAACAGAGAGCTGATA 540
Db      186 ArgIleuProAsnIleuIyrLeuGlyCysGlnArgAspValIleuAsnIleGlyLeuMet 205
QY      541 CAGCAGAAATGGAGTTGGTTATATGTTAAATGCAGCATTAACCTGCAAAAGCTTACCTT 600
Db      206 GlnIleuAsnGlyIleGlyIyrValIleuAsnAlaSerAsnThrCysProIysPheAspPhe 225
QY      601 ATCCCGAGTCTCATTTCCCTGCTGCTGCTGCAATGACAGCTTTTGTGAGAAATTTTG 660
Db      226 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIysIleLeu 245
QY      661 CCGTGTGTTGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCAAATGAGTGT 720
Db      246 ProThrLeuAspIysSerValAspPheIleGluIysAlaIysAlaSerAsnGlyCysVal 265
QY      721 CTAGTCACATGTTTACCTGGATCTCCGCTCCGCTCCGACACATGCTATGCTTACATCATG 780
Db      266 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIyrIleMet 285
QY      781 AAGAGATGAGACATGCTTTAGATGAAGCTTACGATTTGTGAAAGAAAGAAAGAAAGCTTACT 840
Db      286 LysAspGluMetAspMetSerLeuAspGluAlaIyrArgPheValIysGluIysArgProThr 305
QY      841 AATATCTCCAACTTCAATTTTCTGGCCAACTCTGAGACTATGAGAAAGATTAAAGAC 900
Db      306 IleSerProAsnPheAsnPheLeuGlyGlnLeuAspIyrIleIysIleIysAsn 325
QY      901 CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAGCTGCTGACCTGGAGAAAGCAAT 960
Db      326 GlnThrGlyAlaSerGlyProIysSerIyrLeuIysLeuIleuIleuIlySerProAsn 345
QY      961 GAACCTGTCCCTGTGCTCAGAGGGTGGACAGAAAGGAGACCCCTTCAGTCCACC 1020
Db      346 GluProValProAlaValSerGluIysGlyGlnIysSerGluThrProLeuSerProPro 365
QY      1021 TGTGCGCACTGCTGCTCAGAGGAGAGCAAAAGGCGCGTGCATCCGCGCAGGGTG 1080
Db      366 CysAlaAspSerAlaThrSerGluAlaIysGlnArgProValHisProAlaSerVal 385
QY      1081 CCCAGCGTGCAGCGTGCAGCGGTGCTGTTAGAGGACCGCTGTGTACAGCGCTTC 1140
Db      386 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 405
QY      1141 AGTGGGCTGCACTGTGCCAGACAGGGTGGAGACAGCAATAGCTCAAGCCTTCCTC 1200
Db      406 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIyrSleuIyrAspSerPhe 425
QY      1201 TCTTGATATCAATCAGTTTCATATTCAGCGCATGGCAGATCTTACATGGCTTC 1260
Db      426 SerLeuAspIleIysSerValSerIyrSerAlaSerMetAlaIleSerLeuIleGlyPhe 445
QY      1261 TCCATATGAAAGATGCTTGAATACTACAAACCTTCACTCTGATGGAGCCAC 1320
Db      446 SerSerSerGluAspAlaLeuGluIyrIyrIyrProSerThrThrIleuAspGlyThrAsn 465
QY      1321 AAGCTATGCGAGTTCTCCCTGTTGAGAACTATCGGAGAGACATCCGAAACAGTCT 1380
Db      466 LysLeuCysGlnPheSerProValGlnGluLeuSerGlnGlnThrProGluThrSerPro 485
QY      1381 GATAAGAGAGAGAGCCAGCATCCCAAGAGAGCTGACAGCCGAGGCTTACAGACAGCAG 1440
Db      486 AspIysGlnGluAlaSerIleProIysIysLeuGlnThrAlaArgProAspSerGln 505

QY      1441 AGCAGCGAATTCATTCGTTGAGAACAGCAGCAGGACCGCCAGAGATCCCTTTTA 1500
Db      506 SerIyrArgLeuHisSerValAlaArgIleuSerSerSerGlyThrAlaGlnArgSerLeuLeu 525
QY      1501 TCTCCACATTCAGAAATGGAGAGCGTGGAGAGCAATTAACACACAGCTTCTTCCGGC 1560
Db      526 SerProLeuHisArgSerGlySerValGluAspAsnIyrHisThrSerPheLeuPheGly 545
QY      1561 CTTTCACGAGCAGCAGCAGCAGCTCAGAGAGTCTGAGCTGGGCTTAAAGAGCTGGCAC 1620
Db      546 LeuSerThrSerGlnGlnHisLeuThrIyrSerAlaGlyLeuGlyIleuIysGlyIyrHis 565
QY      1621 TCGATATCTTGCCCTCCGAGACCTTACCCCTCCCTGACAGCAGCAGCTGATTTTGC 1680
Db      566 SerAspIleLeuAlaProGlnThrThrSerThrProSerLeuThrSerSerIyrPheAla 585
QY      1681 ACAGATCTCCACATTCATCTGCTGCTGAGGCATTAAGAGGACAGTGCAGATTACTCT 1740
Db      586 ThrGluSerSerHisPheIyrSerAlaSerAlaIleIyrGlyIysSerAlaSerIyrSer 605
QY      1741 GCTTACAGCTCCAGCCAGCTGCTGCTGCTGAGAGCAAGATCTATTTCTGTGCGAGCGG 1800
Db      606 AlatySerCysSerGlnLeuProThrCysGlyAspGlnValIyrSerValAlaArgArg 625
QY      1801 CAGAGCGAAGTACAGAGCTGACTCGCGGAGAGCTGGCATTAAGAGAGAGCCCTTTGAA 1860
Db      626 GlnIysProSerAspArgAlaAspSerArgAspSerIyrPheIysGluIysSerProPheGlu 645
QY      1861 AAGCAGTTTAAACGAGAAAGCTGCAATTTGAGAGAGACATCATGAGAGAAC 1920
Db      646 LysGlnPheIysArgArgSerCysGlnMetGluPheGlyIysSerIleMetSerGluAsn 665
QY      1921 AGGTCACGGAAAGACTGAGGAAAGTGGAGAGTCAAGCTTTTGGGCGACATGAA 1980
Db      666 ArgSerArgGlnGluLeuGlyIysValGlySerGlnSerSerPheSerGlySerMetGlu 685
QY      1981 ATCATTTGAGGTCTCC 1995
Db      686 IleIleGluValSer 690

RESULT 13
US-10-072-012-681
Sequence 681, Application US/10072012
Publication No. US2004003493A1
GENERAL INFORMATION:
APPLICANT: Tchernov, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernusen, Bryan
APPLICANT: Faturajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Beha
APPLICANT: Padigaru, Murajidhara
APPLICANT: Anderson, David W.
APPLICANT: Raselli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Furtak, Karazyna
APPLICANT: Grosse, William M.
APPLICANT: Alsbrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine B.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072, 012
CURRENT FILING DATE: 2002-01-31
```


QY 1621 TCGATATCTTGGCCCCGAGACCTTACCCCTTCTCCGTAACGAGCAGCTGATTTTGGC 1680
 Db 541 Seraspilleuenaalaproginthrsertthproserleuhtsrserlthpyrrheha 560
 QY 1681 ACAAGATCTCACACTTCTACTGCTCAGCCATCTAAGGAGGCGAGTTCCTTACTT 1740
 Db 561 Thrtjuserserhsphetyrsersalaserlailertyglylserlaserltyrser 580
 QY 1741 GCTTACAGCTGAGCGAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCGAGCGCG 1800
 Db 581 AletyrsercysserleueneuProthrCysglYaspGlnValtyrserValArgArgArg 600
 QY 1801 CAGAGCGCAAGTACAGAGCTGACTCGCGCGGAGCTGGCATGAGAGAGAGCCCTTTGAA 1860
 Db 601 GlnlyseroserasparglalaeserlArgargserlthphlsglgluserProphedlu 620
 QY 1861 AACGAGTTTAAAGCGAGAGCTGCCAAATGGAATTGAGAGAGCATGTCAGAGAAC 1920
 Db 621 LysglInPhelyahArgArgsercysglnmetgluphegljuserllemetserGluasn 640
 QY 1921 AGGTCAAGGAGAGAGCTGGGAGAGTGGGAGTCACTAGCTTTTGGGAGAGATGGA 1980
 Db 641 ArgserArggluglInleugllylserValglYserglInserlserpheserGlyserMetGlu 660
 QY 1981 ATCAATGAGGCTCTCC 1995
 Db 661 lletlegluValser 665

RESULT 14

US-10-094-749-2312
 ; Sequence 2312, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: MAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO
 ; APPLICANT: HIO, YURI
 ; APPLICANT: OTSUKA, KAORU
 ; APPLICANT: NAGAI, KEIICHI
 ; APPLICANT: IRIE, RIOTARO
 ; APPLICANT: TAMECHIKA, ICHIRO
 ; APPLICANT: SEKI, NAORIKO
 ; APPLICANT: YOSHIKAWA, TSUTOMU
 ; APPLICANT: OTSUKA, MOTOMYUKI
 ; APPLICANT: NAGAHARI, KENJI
 ; APPLICANT: MASUHO, YASUHIKO
 ; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
 ; FILE REFERENCE: 084335/0160
 ; CURRENT APPLICATION NUMBER: US/10/094,749
 ; PRIOR FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 60/350,435
 ; PRIOR FILING DATE: 2002-01-24
 ; PRIOR APPLICATION NUMBER: JP 2001-328381
 ; NUMBER OF SEQ ID NOS: 3381
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2312
 ; LENGTH: 665
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-094-749-2312

Alignment Scores:

Pred. No.: 1.84e-233 Length: 665
 Score: 3399.00 Matches: 662
 Percent Similarity: 99.70% Conservative: 1
 Best Local Similarity: 99.55% Mismatches: 2

Query Match: 94.02% Indels: 0
 DB: 15 Gaps: 0
 US-10-029-345A-108_COPY_538_2532 (1-1995) x US-10-094-749-2312 (1-665)

QY 1 ATGGCCCATGAGATGATTTGGAATCTCAATTTGTTATCTGAGAGTTGTGCTGTGGGAA 60
 Db 1 MetAlahsglmetllleglYthrglnlleValthrglunrgleuValAlaleuLeuglu 20
 QY 61 AGTGAACGGAAAAAGTGTCTTAATGATAGCCGGCATTTGGTAATTAATTAATTC 120
 Db 21 SerglYthrglulysValleuLeuileasSerlthpProheValglulysrhnThrSer 40
 QY 121 CACTTTTGAAGGCATTAATATTAATCTGCTCCAACTTAATGAAGCAAGTTTGAACAG 180
 Db 41 HstlleuengluAlalleenilleasnlyserlyleuMetlksrghrgleugln 60
 QY 181 GACAAAGTTTAAATTAAGAGCTCATCCAGCATTTGCGGAAACATTAAGTTGACATTGAT 240
 Db 61 AspIysValleuilethrgluleuileglnHlsesrAlalyshlyValAspIleasp 80
 QY 241 TGCAGTGAAGGTTGATTTTCATCAAGCTCCCAAGATGTTGCTCTCTCTCA 300
 Db 81 CysserglulysValValValTyrapglnserSerGlnaspValAlaserleuSerSer 100
 QY 301 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGCTTCAACTCTGTTCACTG 360
 Db 101 AspCysPheleuthrValleuLeugllylsleuglulysSerPheasnSerValHlsleu 120
 QY 361 CTTCAGAGTGGGTTGCTGAGTCTCTGTTGTTTCCCTGGCCTCTGTAAGAAAAATCC 420
 Db 121 LeuAlaglyglyPheleaglnPheSerArgCysPheProglYleuCysglulysSer 140
 QY 421 ACTTGAATCCCACTGATTTCTCAGCTTGTCTTCACTGTTGCAATTTGGCCCAACC 480
 Db 141 ThrleuValProThrCyslleSerglInProCysleuProValAlaasnllleglYProthr 160
 QY 481 CGAATTTCTCCCAATCTTATCTTGCGTCCGAGAGATGCTCTCAACAGAGTGAATA 540
 Db 161 LeuIleleuProasnleuTyrlengllyCysglInArgspValleuLeuulysglleuMet 180
 QY 541 CAGCAAAATGGGATTTGTTATGTTAATGCCAGCATATCTGTGCCAAAGCTTGACTTT 600
 Db 181 GlnGlnasnGlylleglYtyrValleuasnHlsesrhnThrCysProlyspProasphe 200
 QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAAAATTTG 660
 Db 201 lIleProglInserHlsPheleuArgValProvalasnAspSerPheCysglulyslleu 220
 QY 661 CCGTGTGGAACAATCAGTAGATTTCAATTGAGAAAGCAAAAGCTCCCAATGAGATGTT 720
 Db 221 ProThrleuAspIysSerValaspheilleglulysAlalyAsesrhnGlyCysVal 240
 QY 721 CTAGTGACGTTTACGTGGGATCTCCGCTCCGCCACCATGCTATGCTTACATCATG 780
 Db 241 leuValHlsCysleuHlaeglylleSerArgserAlaThrilleAlalyllleMet 260
 QY 781 AAGAGATGACATGCTTAGATGAGCTTACAGATTGGAAGAAAGAAAGCTTACT 840
 Db 261 LysArgMetAspMetSerleuAspGlnAlaTyrapheVallysglulysArgProthr 280
 QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGAGCTATGAGAGAAAGATTAAGAAC 900
 Db 281 lIleSerProasnPhasnPheleuGllyGlnleuLeuaspTyrglulyslyllelyshsn 300
 QY 901 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGCAAGCAAGCAAT 960
 Db 301 GlnthrglyAlaserGlyProlyserlyshleuulshleuGlnlyspProasn 320
 QY 961 GAACCTGTCCTGCTGCTGAGAGGTGAGACAGAAAAAGCAGCGCCCTCATGCAACC 1020
 Db 321 GluproValProAlaValserglulysgllyGlnlyserglInThrProleuSerProPro 340

1021 TGTGCCACTCTGCTACTGAGAGGAGGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1080
1081 CCCAGCGTCCCGCAGCGTGCAGCGTGCCTTTAGAGGAGCCCGCTGTATGAGCGCTC 1140
1081 CCCAGCGTCCCGCAGCGTGCAGCGTGCCTTTAGAGGAGCCCGCTGTATGAGCGCTC 1140
361 ProSerValProSerValGlnProSerLeuGlnuSerProLeuValGlnAlaLeu 380
1141 AGTGGCGTGCAGCTCTGCTCCGAGAGAGGCTGGAAGACAGCAATAGCTCAAGCTTC 1200
381 SerGlyLeuH1LeuSerAlaAspArgLeuGlnuSerH1LeuValGlnuSerPhe 400
1201 TCTCTGATATCAATCAATTCATATTCAGCCAGAGTGCAGATCTTCAATGAGCTTC 1260
401 SerLeuAspLeuLeuSerValSerValSerAlaSerMetAlaAlaSerLeuH1SerGlyPhe 420
1261 TCTCTATCAGAGATGCTTTGGAATATCAATCAATTCAGCTTCATCTGATGGAGCAG 1320
421 SerSerSerGlnuSerAlaLeuGlnuTyTyTyTyProSerThrThrLeuAspGlyThrAsn 440
1321 AACCTATGCGAGTTCCTCCCTGTTTACAGAACTATCGAGAGAGTCCCGAAACCACTCT 1380
441 LysLeuGlnuSerGlnuSerProValGlnuGlnuSerGlnuGlnuProGlnuSerPro 460
1381 GATAGAGAGAGAGAGAGAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
461 AspArgGlnuGlnuAlaSerLeuProValSerLeuGlnuGlnuGlnuGlnuGlnuGlnu 480
1441 AGCAAGAGATGATTCGATTCGATTCAGAACTGAGAGAGAGAGAGAGAGAGAGAGAG 1500
481 SerLysArgLeuH1SerValArgThrSerSerSerGlnuGlnuGlnuGlnuGlnuGlnu 500
1501 TCTCCACTGATGAGAGTGGAGCGTGGAGAGCAATTCAGCAACAGCTTCCTTTGGGC 1560
501 SerProLeuH1AspArgSerGlnuSerValGlnuAspAsnLysH1SerPheLeuPheGly 520
1561 CTTTCCAGCAGCAGCAGCAGCAGCTCAGCAATGCTGCTGGAGCTTAAAGGCTGGCAG 1620
521 LeuSerThrSerGlnuGlnuH1SerLeuH1SerAlaGlnuGlnuGlnuGlnuGlnuGlnu 540
1621 TCGGATATCTTGGCCCCCAGAGCTTACCTTCCCTGACAGCAGAGCTGATTTGGCC 1680
541 SerAspLeuLeuAlaProGlnuThrSerThrProSerLeuThrSerSerThrPheAla 560
1681 ACAGAGTCTCTCAACTTCTGCTGCTGAGCACTTACAGAGAGAGAGAGAGAGTACTCT 1740
561 ThrGlnuSerSerH1AspArgSerAlaSerAlaIleTyGlnuGlnuSerAlaSerTySer 580
1741 GCCTACAGTGCAGCAGCAGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
581 AlaTySerCySerGlnuLeuProH1CySerGlnuAspGlnuAlaTySerValArgArg 600
1801 CAGAGAGAGAGAGAGAGAGTGACTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
601 GlnuLysProSerAspArgAlaAspSerArgArgSerThrPheGlnuGlnuSerProPheGlnu 620
1861 AACGAGTTTAAAGCAGAGAGCTGCAATGGAATTTGGAAGAGCATCATGTCAAGAAC 1920
621 LysGlnuPheLysArgArgSerCySerGlnuMetGlnuPheGlnuGlnuSerGlnuAsn 640
1921 AGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
641 ArgSerArgGlnuGlnuGlnuGlnuGlnuGlnuGlnuGlnuGlnuGlnuGlnuGlnu 660
1981 ATCATGAGAGTCTCC 1995
661 IleIleGlnuValSer 665

APPLICANT: Tchenev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernusen, Bryan
APPLICANT: Patnurajan, Meera
APPLICANT: Shimkete, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Neha
APPLICANT: Padigaru, Mohidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 258
LENGTH: 662
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-258
US-10-029-345A-108_COPY_538_2532 (1-1995) x US-10-072-012-258 (1-662)
Alignment Scores:
Pred. No.: 4,5e-232
Score: 3379.50
Percent Similarity: 99.408
Best Local Similarity: 99.254
Query Match: 93.498
Matches: 662
Conservative: 660
Mismatch: 1
Indels: 3
Gaps: 1
1 ARGGCCCATGAGATGATGGAATGGAATGTTACTGAGAGAGTGGGCTGCTGAGAA 60
1 MetAlaHisGlnuMetIleGlnuThrGlnuIleValThrGlnuArgLeuValAlaLeuGlnu 20
61 AGTGAAGAGAGAGAGAGAGAGTGTGCTAATGATGAGCGGCACTTTGTGGAATACATATCC 120
21 SerGlyThrGlnuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:26 ; Search time 25.4447 Seconds

(without alignments)
8095.510 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 3615
Sequence: 1 atggcccatgagatgcttg.....tggaatcatgagctcc 1995

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10029345/runat_21062004_122816_4233/app_query.fasta_1.3278
-DB=Issued Patents AA -OPMT=factan -SUFFIX=rai -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029345_@CIGN_1_1_45 @runat_21062004_122816_4233 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/58_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/58_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3406	94.2	665	US-09-816-494-2	Sequence 2, App1
2	654.5	18.1	170	US-09-544-716-14	Sequence 14, App1
3	654.5	18.1	170	US-09-557-921-15	Sequence 15, App1
4	654.5	18.1	170	US-09-564-357-17	Sequence 17, App1
5	654.5	18.1	170	US-09-619-380-16	Sequence 16, App1
6	469	13.0	482	US-09-557-921-2	Sequence 2, App1
7	455	12.6	394	US-09-702-705-805	Sequence 805, App
8	455	12.6	394	US-09-736-457-805	Sequence 805, App
9	455	12.6	394	US-09-614-124B-805	Sequence 805, App
10	455	12.6	394	US-09-671-325-805	Sequence 805, App
11	455	12.6	394	US-09-589-184-805	Sequence 805, App
12	447	12.4	394	US-08-530-290-23	Sequence 23, App1

13	447	12.4	394	US-09-702-705-827	Sequence 827, App
14	447	12.4	394	US-09-736-457-827	Sequence 827, App
15	447	12.4	394	US-09-614-124B-827	Sequence 827, App
16	447	12.4	394	US-09-671-325-827	Sequence 827, App
17	447	12.4	394	US-09-589-184-827	Sequence 827, App
18	439.5	12.2	395	US-08-990-379-5	Sequence 5, App1
19	433	12.0	367	US-08-990-379-6	Sequence 6, App1
20	425	11.8	314	US-09-371-671B-11	Sequence 11, App1
21	420	11.6	367	US-08-530-290-24	Sequence 24, App1
22	409.5	11.3	313	US-08-990-379-7	Sequence 7, App1
23	409	11.3	314	US-09-164-193-22	Sequence 22, App1
24	409	11.3	314	US-09-221-448A-22	Sequence 22, App1
25	393.5	10.9	393	US-08-990-379-4	Sequence 4, App1
26	377	10.4	302	US-09-702-705-806	Sequence 806, App
27	377	10.4	302	US-09-736-457-806	Sequence 806, App
28	377	10.4	302	US-09-614-124B-806	Sequence 806, App
29	377	10.4	302	US-09-671-325-806	Sequence 806, App
30	377	10.4	302	US-09-589-184-806	Sequence 806, App
31	366	10.1	397	US-08-990-379-8	Sequence 8, App1
32	346.5	9.6	168	US-09-544-716-13	Sequence 13, App1
33	346.5	9.6	168	US-09-557-921-13	Sequence 13, App1
34	346.5	9.6	168	US-09-564-357-16	Sequence 16, App1
35	346.5	9.6	168	US-09-619-380-15	Sequence 15, App1
36	338.5	9.4	170	US-09-544-716-12	Sequence 12, App1
37	338.5	9.4	170	US-09-557-921-12	Sequence 12, App1
38	338.5	9.4	170	US-09-564-357-15	Sequence 15, App1
39	338.5	9.4	170	US-09-619-380-14	Sequence 14, App1
40	319	8.8	169	US-09-544-716-16	Sequence 16, App1
41	319	8.8	169	US-09-557-921-17	Sequence 17, App1
42	319	8.8	169	US-09-564-357-19	Sequence 19, App1
43	319	8.8	169	US-09-619-380-18	Sequence 18, App1
44	312	8.6	168	US-09-544-716-15	Sequence 15, App1
45	312	8.6	168	US-09-557-921-16	Sequence 16, App1

ALIGNMENTS

RESULT 1
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCES: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816, 494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191, 858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-816-494-2

Alignment Scores:
Pred. No.: 7.74e-300
Score: 3406.00
Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 94.22%
DB: 4
Gaps: 0

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-816-494-2 (1-665)

QY	1	ATGGCCCATGAGATGATGGAAGTCTGATGAGAGCTGCTGCTGCGGAA	60
DB	1	MetAlhIeGlWetIleGlyThrGlnIleValThhGluArgIleValAlaLeuLgIu	20
QY	61	AGTGAACGGAAGAAAGTGTGCTAATGATGATGACCGGCATTGTGGAATACATATCC	120

Db 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
QY 121 CACATTTTGGAGCCCAATTAATCAACTGCTCCAGCTTATGAGACCGAAGTTGGCAACAG 180
Db 41 HisIleuGlnAlaIleAsnIleAsnSerIleuMetCysArgGluGln 60
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCAATTCAGGAAACATTAAGTTGACATTGAT 240
Db 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY 241 TGAGTCGAGAGAGTTGATGTTAGATCAAGCTCCCAAGATGTTGCTCTCTCTTTCA 300
Db 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspAlaIleSerLeuSer 100
QY 301 GACTGTTTCTCACTGTAATCTTGGGTAACCTGGAAGAGCTTCAACTGTGTACCTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisIleu 120
QY 361 CTGGAGGTGGGTTTGGTGAAGTTCTCGTGTGTTTCCCTGGGCTCTGGAAGGAAATCC 420
Db 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
QY 421 ACTTAAGTCCCTACTGCAATTTCTCAGCCTTGAATCTGTTGCCAATTTGGGCCAAC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTTATCTGGCTGCCAGAGATGTCCTCAACAGAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgPheValLeuAsnLysGluLeuMet 180
QY 541 CAGCAATAGGATTTGTTATGTTAAATGCAAGCTATACCTGTCCAAAGCTGACTT 600
Db 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGTAATGACAGCTTTTGTGAGAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
QY 661 CCGTGGTGGCAAAATCAATGATTTGATGAGAAAGCAAAAGCTCCCAATGATGTTT 720
Db 221 ProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnIleCysVal 240
QY 721 CTAGTGCACGTGTAGCTGGGATCTCCGCTCCGCCACATGCTATGCTCACTCATG 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
QY 781 AAGAGATGACATGTTCTTAAATGATGAGCTTACAGATTTGGAAGAAAGAAAGCTACT 840
Db 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
QY 841 ATATCTCAAACTTCAATTTTCCGGCCCAATCTCCGACATATGAGAAAGATTAAGAAC 900
Db 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuAspLysTyrGluLysIleLysAsn 300
QY 901 CAGACGTGAGCATCAGGCGCAAAAGCAAACTCAAGCTGCTGACCTCGGAGAGCCCAAT 960
Db 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuMetHisLeuLysProAsn 320
QY 961 GAACCTGTCTCTGTCTCAGAGGTTGACAGAAAGCAAGCCCTCATGCTCACCC 1020
Db 321 GluProValProAlaValSerGluGlyGlnLysSerGlnThrProLeuSerProPro 340
QY 1021 TGTGCGCATCTGCTACTCTCAGAGCAGAGCAAAAGCCCTGATCCCGCCAGCTG 1080
Db 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 1080
QY 1081 CCCAGCTGCCAGCTGACGCTGCTGCTGTTAGAGAGCAAGCCCTGCTGTAACAGCTGCT 1140
Db 361 ProSerValProSerValGlnProSerLeuGlnAspSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGCTGACCTGTCCTCCGAGACAGCTGGAAGACAGCAATTAAGCTCAACGTTCTTC 1200

Db 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
QY 1201 TCTTGATATTCAAATCAGTTTCAATTCATGACGACAGACAGATCTTATCATGCTTCC 1260
Db 401 SerLeuAspIleLysSerValSerLysSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
QY 1261 TCCATCATGAGAAGATGCTTTGGAAATCTAACAACTTCCACTCATCTGTGATGGACCAAC 1320
Db 421 SerSerSerGlnAspAlaLeuGluLysTyrLysProSerThrThrLeuAspGlyThrAsn 440
QY 1321 AAGCTATGCGCACTTCTCCCTGTTCAAGAACTATCGAGAGCACTCCGAGGCTTTCAGACAGCCAG 1440
Db 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGlnGlnThrProGluThrSerPro 460
QY 1381 GATTAAGAGAGAGCAGATCCCAAAACCTGCAACCGCCAGGCTTTCAGACAGCCAG 1440
Db 461 AspLysGlnGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
QY 1441 AGCAAGCATTTGCAATTTGTCAGAACCCAGACAGAGAGAGAGCCAGGCTTCTTTTA 1500
Db 481 SerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArgSerLeuLeu 500
QY 1501 TCTCCATGATGGAAGTGGAGCGTGGAGCAATTAACACACAGCTTCTTTGGGC 1560
Db 501 SerProLeuHisArgSerGlySerValGlnAspAsnTyrHisSerPheLeuPheGly 520
QY 1561 CTTTCCACAGCCAGACAGCACTCAAGAGTGTGCTGCTGCTGCTTAAAGGCTGGCAC 1620
Db 521 LeuSerThrSerGlnHisIleLeuThrLysSerAlaGlyLeuGlyLeuLysGlyThrPhe 540
QY 1621 TCGGATATTTGGCCCCCAGACCTCTACCCCTCCCTGACACAGCAAGCTGTATTTGGC 1680
Db 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerIlePyrPheAla 560
QY 1681 ACAGAGTCTCACTTTACTCTGCTCAGCCATCTACCGAGAGAGAGCTGACAGTTACT 1740
Db 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysAlaSerLysSer 580
QY 1741 GCTTACAGCTCAGCCAGCTGCCCCACTTTCGAGAGACCAAGTCTATTTCTTGGACGCG 1800
Db 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAsnGlnValTyrSerValArgArg 600
QY 1801 CAGAGCCAGATGACAGAGCTGCTGCGCGGAGAGCTGGCATGAAGAGAGCCCTTTGAA 1860
Db 601 GlnLysProSerAspArgAlaAspSerArgArgSerIleGlnGluSerProPheGlu 620
QY 1861 AAGCAATTTTAAACGCAAGAGCTGCCAAATGAAATTTGAGAGAGATCATCTCAGAAC 1920
Db 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyLysIleMetSerGluAsn 640
QY 1921 AGGTCAAGGAGAGAGCTGGGGAATGTGGCGAGTCACTTACCTTTTCCGAGACATGA 1980
Db 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
QY 1981 ATGATTTGAGCTTCC 1995
Db 661 IleIleGluValSer 665

RESULT 2
US-09-544-716-14
Sequence 14, Application US/09544716
Patent No. 6492157
GENERAL INFORMATION:
APPLICANT: Lucite, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-9 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: DSP-9 DUAL-SPECIFICITY PHOSPHATASE
CURRENT APPLICATION NUMBER: US/09/544, 716
CURRENT FILING DATE: 2000-04-10
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 170


```
TYPE: PRT
ORGANISM: Homo sapiens
US-09-544-716-14

Alignment Scores:
Pred. No.: 6,75e-51 Length: 170
Score: 654.50 Matches: 121
Percent Similarity: 88.62% Conservative: 27
Best Local Similarity: 72.46% Mismatches: 18
Query Match: 18.11% Indels: 1
DB: 4 Gaps: 1

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-544-716-14 (1-170)
QY 400 GGCCTCTGTGAAGAAA---TCACCTTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTA 456
DB 1 GlyLeuCysGlnGlyLysProAlaAlaLeuLeuProMetSerLeuSerGlnProCysLeu 20
QY 457 CCTGTGGCCAAACATTTGGGCGCAACCCGAATCTTCCCAATCTTATCTTGCTGGCCGCGA 516
DB 21 ProValProSerValGlyLeuThrArgGlyLeuProHileuProHileuGlySerGlnLys 40
QY 517 GATGTCTCAACAAGAGCTGATACAGAGAAATGGATGGATTGATTGTTAAATGGCAGC 576
DB 41 AspValLeuAsnLysAspLeuMetThrGlnAsnGlyLeuSerValLeuAsnAlaSer 60
QY 577 TATACCTGTCAAAGCCTGACTTATCCCGAGTCTCATTTCTGCGCTGCTGTAAT 636
DB 61 AsnSerCysProLysProAspPheIleCysGlnSerArgPheMetArgValProIleAsn 80
QY 637 GACAGCTTTTGTGAGAAAATTTTGGCGGTGGAGCAAAATCAGTAATTTCAATTGAGAA 696
DB 81 AspAsnTyrCysGlnLysLeuLeuProTrpLeuAspLysSerIleGluPheIleAspLys 100
QY 697 GCAAAAGCCTCCAAATGATGTGTTCTAGTGCACATGTTTACGGGATCTCCCGCTCGCC 756
DB 101 AlAlaLysLeuSerSerCysGlnValIleValHisCysLeuAlaGlyIleSerArgSerAla 120
QY 757 ACCATGCTATGCGCTTACATCATGAGAGAGATGATGCTTTAGTAGAAGCTTACAGA 816
DB 121 ThrIleAlaIleAlaTyrIleMetLysThrMetGlyMetSerSerAspAlaTyrArg 140
QY 817 TTTGTGAAAAGAAAAGAACTTACTATATCTCCAAATCTCAATTTTCTGGGCGCAATCTCG 876
DB 141 PheValLysAspArgArgProSerIleSerProAsnPheAsnPheLeuGlyGlnLeuLeu 160
QY 877 GACTATGAGAAGATTAG 897
DB 161 GluTyrGluArgThrLeuLys 167

RESULT 3
US-09-557-921-15
; Sequence 15, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: LucHE, Ralf M.
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125, 416
; CURRENT APPLICATION NUMBER: US/09/557, 921
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-921-15

Alignment Scores:
Pred. No.: 6,75e-51 Length: 170
Score: 654.50 Matches: 121
Percent Similarity: 88.62% Conservative: 27
```

```
Best Local Similarity: 72.46% Mismatches: 18
Query Match: 18.11% Indels: 1
DB: 4 Gaps: 1

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-557-921-15 (1-170)
QY 400 GGCCTCTGTGAAGAAA---TCACCTTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTA 456
DB 1 GlyLeuCysGlnGlyLysProAlaAlaLeuLeuProMetSerLeuSerGlnProCysLeu 20
QY 457 CCTGTGGCCAAACATTTGGGCGCAACCCGAATCTTCCCAATCTTATCTTGCTGGCCGCGA 516
DB 21 ProValProSerValGlyLeuThrArgGlyLeuProHileuProHileuGlySerGlnLys 40
QY 517 GATGTCTCAACAAGAGCTGATACAGAGAAATGGATGGATTGATTGTTAAATGGCAGC 576
DB 41 AspValLeuAsnLysAspLeuMetThrGlnAsnGlyLeuSerValLeuAsnAlaSer 60
QY 577 TATACCTGTCAAAGCCTGACTTATCCCGAGTCTCATTTCTGCGCTGCTGTAAT 636
DB 61 AsnSerCysProLysProAspPheIleCysGlnSerArgPheMetArgValProIleAsn 80
QY 637 GACAGCTTTTGTGAGAAAATTTTGGCGGTGGAGCAAAATCAGTAATTTCAATTGAGAA 696
DB 81 AspAsnTyrCysGlnLysLeuLeuProTrpLeuAspLysSerIleGluPheIleAspLys 100
QY 697 GCAAAAGCCTCCAAATGATGTGTTCTAGTGCACATGTTTACGGGATCTCCCGCTCGCC 756
DB 101 AlAlaLysLeuSerSerCysGlnValIleValHisCysLeuAlaGlyIleSerArgSerAla 120
QY 757 ACCATGCTATGCGCTTACATCATGAGAGAGATGATGCTTTAGTAGAAGCTTACAGA 816
DB 121 ThrIleAlaIleAlaTyrIleMetLysThrMetGlyMetSerSerAspAlaTyrArg 140
QY 817 TTTGTGAAAAGAAAAGAACTTACTATATCTCCAAATCTCAATTTTCTGGGCGCAATCTCG 876
DB 141 PheValLysAspArgArgProSerIleSerProAsnPheAsnPheLeuGlyGlnLeuLeu 160
QY 877 GACTATGAGAAGATTAG 897
DB 161 GluTyrGluArgThrLeuLys 167

RESULT 4
US-09-564-357-17
; Sequence 17, Application US/09564357
; Patent No. 6645753
; GENERAL INFORMATION:
; APPLICANT: LucHE, Ralf M.
; TITLE OF INVENTION: DSP-5 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125, 413
; CURRENT APPLICATION NUMBER: US/09/564, 357
; CURRENT FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-564-357-17

Alignment Scores:
Pred. No.: 6,75e-51 Length: 170
Score: 654.50 Matches: 121
Percent Similarity: 88.62% Conservative: 27
Best Local Similarity: 72.46% Mismatches: 18
Query Match: 18.11% Indels: 1
DB: 4 Gaps: 1

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-564-357-17 (1-170)
QY 400 GGCCTCTGTGAAGAAA---TCACCTTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTA 456
DB 1 GlyLeuCysGlnGlyLysProAlaAlaLeuLeuProMetSerLeuSerGlnProCysLeu 20
QY 457 CCTGTGGCCAAACATTTGGGCGCAACCCGAATCTTCCCAATCTTATCTTGCTGGCCGCGA 516
DB 21 ProValProSerValGlyLeuThrArgGlyLeuProHileuProHileuGlySerGlnLys 40
QY 517 GATGTCTCAACAAGAGCTGATACAGAGAAATGGATGGATTGATTGTTAAATGGCAGC 576
DB 41 AspValLeuAsnLysAspLeuMetThrGlnAsnGlyLeuSerValLeuAsnAlaSer 60
QY 577 TATACCTGTCAAAGCCTGACTTATCCCGAGTCTCATTTCTGCGCTGCTGTAAT 636
DB 61 AsnSerCysProLysProAspPheIleCysGlnSerArgPheMetArgValProIleAsn 80
QY 637 GACAGCTTTTGTGAGAAAATTTTGGCGGTGGAGCAAAATCAGTAATTTCAATTGAGAA 696
DB 81 AspAsnTyrCysGlnLysLeuLeuProTrpLeuAspLysSerIleGluPheIleAspLys 100
QY 697 GCAAAAGCCTCCAAATGATGTGTTCTAGTGCACATGTTTACGGGATCTCCCGCTCGCC 756
DB 101 AlAlaLysLeuSerSerCysGlnValIleValHisCysLeuAlaGlyIleSerArgSerAla 120
QY 757 ACCATGCTATGCGCTTACATCATGAGAGAGATGATGCTTTAGTAGAAGCTTACAGA 816
DB 121 ThrIleAlaIleAlaTyrIleMetLysThrMetGlyMetSerSerAspAlaTyrArg 140
QY 817 TTTGTGAAAAGAAAAGAACTTACTATATCTCCAAATCTCAATTTTCTGGGCGCAATCTCG 876
DB 141 PheValLysAspArgArgProSerIleSerProAsnPheAsnPheLeuGlyGlnLeuLeu 160
QY 877 GACTATGAGAAGATTAG 897
DB 161 GluTyrGluArgThrLeuLys 167
```

```

Db      1 GlyLeuCysGluGlyLysProAlaLeuLeuProMetSerLeuSerGlnProCysLeu 20
QY      457 CCTGTTCGCAACATTGGGCCAACCCGAATTTCTTCCAAATCTTTATCTTGGCTGCCACGCA 516
Db      21 ProValProSerValGlyLeuThrArgGlnLeuProHsLeuGlySerGlnLys 40
QY      517 GATGTCTTCACAAAGAGCTGATACAGCAAGATGGAGTTGGTTATGTGTAAATGCGCAGC 576
Db      41 AspValLeuAsnLysAspLeuMetThrGlnAsnGlyLeuSerTyrValLeuAsnAlaSer 60
QY      577 TATACCTGTCGAAGCTGATCTTATTCCTCCAGATCTCATCTCCGGCTGGCTGGAT 636
Db      61 AsnSerCysProLysProAspPheIleCysGlnSerArgPheMetArgValProIleAsn 80
QY      637 GACAGCTTTTGGAGAAATTTTGGCGGTGGATGACAAATCAGTATGATGAACTTACAGAA 696
Db      81 AspAsnTyrCysGlnLysLeuLeuProIlePheAspLysSerIleGlnPheIleAspLys 100
QY      697 GCAAAAGCTCCAAATGATGTGTTCTAGTGCACGTGTTAGCTGGAGATCCCGCTCCGCC 756
Db      101 AlalysLeuSerSerCysGlnValIleValHisCysLeuAlaGlyIleSerArgSerAla 120
QY      757 ACCATCGCTATCGCTCATCATGAGAGAGATGACATGTCTTATGATGAACTTACAGCA 816
Db      121 ThrIleAlaIleAlaTyrIleMetLysThrMetGlyMetSerSerAspAlaTyrArg 140
QY      817 TTGTGAAAGAAAGAAAGACCTACTATATCTTCCAAATTTCTTGGCGCACTCTCG 876
Db      141 PheValLysAspArgPheProSerIleSerProAsnPheAsnPheLeuGlyGlnLeuLeu 160
QY      877 GACTATGAGAAAGATTAAG 897
Db      161 GlutyrGlnArgThrLeuLys 167

```

RESULT 5
US-09-619-380-16

```

/ Sequence 16, Application US/09619380
/ Patent No. 6649391
/ GENERAL INFORMATION:
/ APPLICANT: Lucbe, Ralf M.
/ TITLE OF INVENTION: DSP-11 DUAL SPECIFICITY PHOSPHATASE
/ FILE REFERENCE: 200125.418
/ CURRENT APPLICATION NUMBER: US/09/619,380
/ CURRENT FILING DATE: 2000-07-19
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 16
/ LENGTH: 170
/ TYPE: PRN
/ ORGANISM: Homo sapiens
US-09-619-380-16

```

Alignment Scores:
Pred. No.: 6,75e-51
Score: 654.50
Percent Similarity: 88.62%
Best Local Similarity: 72.46%
Query Match: 18.11%
Matches: 121
Conservative: 27
Mismatches: 18
Indels: 1
Gaps: 1

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-619-380-16 (1-170)

```

QY      400 GGCCCTGTGAAAGAAA--TCACCTTAGTCCCTACGCACTTCTTCCAGCTTGTCTTA 456
Db      1 GlyLeuCysGluGlyLysProAlaLeuLeuProMetSerLeuSerGlnProCysLeu 456
QY      457 CCTGTTCGCAACATTGGGCCAACCCGAATTTCTTCCAAATCTTTATCTTGGCTGCCACGCA 516
Db      21 ProValProSerValGlyLeuThrArgGlnLeuProHsLeuGlySerGlnLys 40
QY      517 GATGTCTTCACAAAGAGCTGATACAGCAAGATGGAGTTGGTTATGTGTAAATGCGCAGC 576

```

```

Db      41 AspValLeuAsnLysAspLeuMetThrGlnAsnGlyIleSerTyrValLeuAsnAlaSer 60
QY      577 TATACCTGTCGAAGCTGATCTTATTCCTCCAGATCTCATCTCCGGCTGGCTGGAT 636
Db      61 AsnSerCysProLysProAspPheIleCysGlnSerArgPheMetArgValProIleAsn 80
QY      637 GACAGCTTTTGGAGAAATTTTGGCGGTGGATGACAAATCAGTATGATGAACTTACAGAA 696
Db      81 AspAsnTyrCysGlnLysLeuLeuProIlePheAspLysSerIleGlnPheIleAspLys 100
QY      697 GCAAAAGCTCCAAATGATGTGTTCTAGTGCACGTGTTAGCTGGAGATCCCGCTCCGCC 756
Db      101 AlalysLeuSerSerCysGlnValIleValHisCysLeuAlaGlyIleSerArgSerAla 120
QY      757 ACCATCGCTATCGCTCATCATGAGAGATGACATGTCTTATGATGAACTTACAGCA 816
Db      121 ThrIleAlaIleAlaTyrIleMetLysThrMetGlyMetSerSerAspAlaTyrArg 140
QY      817 TTGTGAAAGAAAGAAAGACCTACTATATCTTCCAAATTTCTTGGCGCACTCTCG 876
Db      141 PheValLysAspArgPheProSerIleSerProAsnPheAsnPheLeuGlyGlnLeuLeu 160
QY      877 GACTATGAGAAAGATTAAG 897
Db      161 GlutyrGlnArgThrLeuLys 167

```

RESULT 6
US-09-557-921-2

```

/ Sequence 2, Application US/09557921
/ Patent No. 6551810
/ GENERAL INFORMATION:
/ APPLICANT: Lucbe, Ralf M.
/ TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
/ FILE REFERENCE: 200125.416
/ CURRENT APPLICATION NUMBER: US/09/557,921
/ CURRENT FILING DATE: 2000-04-20
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 482
/ TYPE: PRN
/ ORGANISM: Homo sapiens
US-09-557-921-2

```

Alignment Scores:
Pred. No.: 7.49e-34
Score: 469.00
Percent Similarity: 57.72%
Best Local Similarity: 35.91%
Query Match: 12.97%
Matches: 107
Conservative: 65
Mismatches: 96
Indels: 30
Gaps: 7

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-557-921-2 (1-482)

```

QY      79 CTGCTAATTGATAGCCGCGCATTTGTGGAATACATACATCCCAATTTGGAAGCATT 138
Db      173 ValIleIleAspCysArgProPheMetClnLysAsnLysSerHisIleGlnAlaVal 192
QY      139 AATATCAACTGCTCC--AAGCTTATGACGAGAGGTTGCAACAGCAAAAGGTTAATT 195
Db      193 HisIleAsnCysAlaAspLysIleSerArgArgLysGlnGlnGlnLysIleThrVal 212
QY      196 ACAGAGCTATCCAG--CATTCAGCAAGACATTAAGTTGACATGATTCAGTCAGAG 252
Db      213 LeuAspLeuIleSerCysArgGluGlyLysAspSerPheLysArgIlePheSerLysGln 232
QY      253 GTTGAAGTTTACATCAAACTCCCAAGAGTTGCTCTCTCTTTCAGACTGTTTCTC 312
Db      233 IleIleValTyrAspGlnMetThrAsnGlnProSerArgValMetProSerGlnProLeu 252
QY      313 ACTTACTTCTGGGTAACTGAGGAAGACCTTCACTGTGTCACCTGCTTCAGGTGG 372

```

```

Db      253 HsE1LeValLeuGlnSerLeuLysArgGluGlyLeuGlnProLeuValLeuLysGlyGly 272
Qy      373 TTGGCTGAGTCTCTGCTGTTTCCCTGGCGCTCTGTGACAGAAAA----- 417
Db      273 LeuSerSerPheLysGlnAsnH1sGlnAsnLeuCySerAsnSerLeuGlnLeuGln 292
Qy      418 -----TCCACTAGTCCCTACCTGATTTCT 444
Db      293 CyArgGluValGlyGlyAlaSerAlaLysSerLeuProGlnProIle--- 311
Qy      445 CAGCTTGCTTACTCT--GTGGCAACATTGGGCAACCGAAATCTTCCCAATCTTAT 501
Db      312 ---ProThrThrProArgPrlLeuAsnAlaGlnLeuThrProIleuProPheLeuPhe 330
Qy      502 CTGGGCTCCAGAGATGTCTTCACACAGAGACTGATACAGAAATGGAGTTGGTAT 561
Db      331 LeuGlnAsnGlnGlnAsnAlaAsnLeuAsnThrMetGlnArgLeuAsnIleGlyTyr 350
Qy      562 GTGTAAATGCGC-----TATACCTGCAAGAGCTGACTTATCCCC 606
Db      351 ValIleAsnValThrThrIleuProLeuTyrH1sTyrGlnLysGlyLeuPhe----- 368
Qy      607 GAGTCTCAATTCCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTGGCGTGG 666
Db      369 -----AsnTyrLysArgLeuProAlaThrAspSerAsnLysGlnAsnLeuArgGlnTyr 386
Qy      667 TTGGACAAATCAAGTATTCATTGAGAAAGCAAAAGCTCCAAATGAGATGTGTCTATG 726
Db      387 PheGlnGlnLysPheGlnPheIleGlnLysAlaH1sGlnLysGlyLysGlyLeuIle 406
Qy      727 CACTGTATGCTGGAGATCTCCGCTCGCCAGCAGCAGCTGATGCGCTATGCAATGAGAGAG 786
Db      407 HsCyGlnLysLysValSerArgSerAlaThrIleAlaIleAlaTyrLeuMetLysH1s 426
Qy      787 ATGACATGTCTTATGATGAGCTTACAGATTTGTGAAAGAAAAGACTTATATCT 846
Db      427 ThrArgMetThrMetThrAspAlaTyrLysPheValLysGlyLysArgProIleIleSer 446
Qy      847 CCAACTTCATTTTCTGGGCGCAACTCTGAGACTATGAGAAAGATTAAAGAC 900
Db      447 ProAsnLeuAsnPheMetGlyGlnLeuGlnLysPheGlnLysPheLeuAsn 464

RESULT 7
US-09-702-705-805
; Sequence 805, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvik, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-702-705-805

Alignment Scores:
Pctd. No.: 1,25e-32 Length: 394
Score: 455.00 Matches: 122
Percent Similarity: 49.87% Conservative: 75

```

```

Best Local Similarity: 30.89% Mismatches: 143
Query Match: 12.59% Indels: 55
Db: 4 Gaps: 13

US-10-029-345a-108_copy_538_2532 (1-1995) x US-09-702-705-805 (1-394)

Qy      45 GGTGCTCTGCTGCAAGATGAAACGGA-----AAAGT 77
Db      26 GlyIleAlaGlyLysSerIleSerH1sGlnThrLeuGlyLeuProSerGlyLysCy 45
Qy      78 GCTGTAATGATGACCGCCATTTGTGAAATCAATACATCCCAATTTGAGACCAT 137
Db      45 sLeuLeuAsnCyArgProPheLeuAlaH1sSerAlaGlyTyrIleLeuLysSerVa 65
Qy      138 TATATCACTGTCTCAACCTATATGAAAGGAGTTGCAACAGACAAAGTTAATAC 197
Db      65 lAsnValArgCyAsnThrIleValArgArgArg--AlaLysGlySerLysSerLeuG 84
Qy      198 AGAGCTCATC-----CAGCATTCAGCGAAACATAAGTTGACATTGACATGACAGAA 251
Db      84 uGlnIleuProAlaGlnGlnLysValArgAlaArgLysGlyLeuTyrSerAl 104
Qy      252 GGTGTAGTTTACATCAAGCTCCCAAGATGTGCTCTCTTCAAGCTGTTTCT 311
Db      104 aValIleValTyrAspGlnArgSerProArgLysSerLeuArgGlnLysSerThrVa 124
Qy      312 CACTGACTTCTGGGTAACCTGAGAAAGCTTC-----AACTGTTCACCTGCTTGC 365
Db      124 lSerLeuValAlaGlnAlaLeuArgArgAsnAlaGlnArgThrAspIleCyLeuLeu 144
Qy      366 AGGTGGATTGTGCTGAGTCTCTGTTTCCCTGCGCTCTGTGAAAGAAATTCACCT 425
Db      144 sGlyLysTyrGlnArgPheSerSerGlnTyrProGlnPheCySerLysThrLysAla 164
Qy      426 A-----GTCCCTACCTGCAATTTCTACGCT-----TGGTT 455
Db      164 lAlaAlaIleProProProValProProSerAlaThrGlnProLeuAsnLysCySe 184
Qy      456 ACCTGTGCC-----AACATTGGGCGCAACCCCAATTTCTCCCAATCTT 500
Db      184 rSerCyGlyThrProLeuH1sAsnGlnLysGlnProValGlnIleLeuProPheLeu 204
Qy      501 TCTTGGCTGCGACGAGATGTCTCAACAGAGCTGATACAGCAGATGGATGGTTA 560
Db      204 rLeuGlnSerAlaTyrH1sAlaAlaArgArgPheMetLeuAsnLysIleThrAl 224
Qy      561 TGTGTTAAATGCGCAGCTATATCTGTCAAAGCTGATTTATCCCGAGTCTATTTCT 620
Db      224 aLeuLeuAsnValSerSerAspCySerProAsn--H1sPheGlnGlyH1sTyrGlnTyr 243
Qy      621 GCGTGTCTGTGATGACAGCTTTTGTGAAAAATTTGCGGTGTGACAAATCACT 680
Db      243 sCyIleProValGlnAsnAsnH1sLysAlaAsnIleSerSerTyrPheMetGlnAla 263
Qy      681 AAGTTTCATTGAGAAAGCAAAAGCTCCAAATGAGTGTGTATGACATGTTACTGTG 740
Db      263 eGlnTyrIleAsnPalValLysPheCyArgGlnArgValLeuValH1sCySerAlaG 283
Qy      741 GATTCCTCCGCTCGCACCATGCTATCCCTATCATATGAAAGAGATGAGACATGCTTT 800
Db      283 yIleSerArgSerAlaThrIleCyLeuAlaTyrLeuMetMetLysLysAlaArgLe 303
Qy      801 AGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTATATCTTCCAACTTCAATT 860
Db      303 uGlnGlnAlaPheGlnPheValLysGlnArgArgSerIleIleSerProAsnPheSerP 323
Qy      861 TCTGGGCAACTCTGAGTATGAGAAAGAAATTGAAGACAGACAGACTGACAGGCGC 920
Db      323 eMetGlyGlnLeuLeuGlnPheGlnSerGlnValLeuAlaThrSerCyAlaAlaGln 343
Qy      921 AAAGACAAATCTCAAGCTCTGCACTGAGAGAGCAAAATGAACCTGCTCTGCTTC 980
Db      343 aLysSer-----ProSerGlyProLeu----- 350

```

```

OY 981 AGAGGTGACAGAAAAGCGAGCGCCCTGATCCAGCTGTGGCAGCTGTGTAATTC 1040
Db 351 -----GlyGluArgGlyLeuThrProAlaThrPro-----ThrSe 362
OY 1041 AGAGGACAGAGAGCAAAAGCCCGTCATCCGCCAGCGTCCCGCAGCTGCA 1100
Db 362 rGlnPheValIpheserPheProVal---SerValGlyValHisSerAlaProSerSerle 381
OY 1101 GCCGTGCTGTAGAGAGACGCCCGCTGTACAGCGCTCAGT 1143
Db 381 uProTyLeu-----HisSerProIleThrThrSerProSer 393

```

RESULT 8

```

US-09-736-457-805
/ Sequence 805, Application US/09736457
/ Patent No. 6509448
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darick
/ APPLICANT: Recter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ APPLICANT: Wang, Aijun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: DIAGNOSIS OF LUNG CANCER
/ CURRENT APPLICATION NUMBER: US/09/736,457
/ NUMBER OF SEQ ID NOS: 2000-12-13
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 805
/ LENGTH: 394
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-09-736-457-805

```

Alignment Scores:

```

Pred. No.: 1,25e-32 Length: 394
Score: 455.00 Matches: 122
Percent Similarity: 49.87% Conservative: 75
Best Local Similarity: 30.89% Mismatches: 143
Query Match: 12.59% Indels: 55
Gaps: 13

```

US-10-029-345a-108_copy_538_2532 (1-1995) x US-09-736-457-805 (1-394)

```

OY 45 GGTGGCTCTGCTGGAAGTGAACGAA-----AAAGT 77
Db 26 G|Y|G|A|G|Y|S|e|L|S|e|H|G|Y|T|H|e|U|G|Y|e|U|P|r|o|S|e|r|G|Y|e|U|Y|S|C| 45
OY 78 GCGTCATATGATAGCGGCCCATTTGTGAAATACAAATCCACATTTTGGAGGCCAT 137
Db 45 s|e|u|e|u|e|u|a|b|C|y|A|r|g|P|r|o|H|e|L|a|H|S|e|r|A|G|Y|T|T|L|e|u|G|Y|S|e|r|V|A 65
OY 138 TAATATCACTGCTCCAGCTTATAGACCGAAGGTTGCAACAGACAAAGTGTATATAC 197
Db 65 l|a|n|V|a|l|g|y|S|a|n|T|H|I|e|V|a|l|A|r|g|A|r|g|A|g|y|---A|A|Y|G|Y|S|e|r|V|a|l|S|e|r|L|e|U|G| 84
OY 198 AGAGCTATC-----CAGCATTCAGCGAAACATTAAGTTGACATGATTCGACATGCAAGTCAAG 251
Db 84 u|G|I|n|L|e|u|P|r|o|A|G|I|n|G|I|n|U|V|a|l|A|r|g|A|a|r|g|L|e|u|A|g|S|e|r|G|Y|L|e|U|Y|S|e|r|A| 104
OY 252 GGTGTAGTTATTCATCAAGCTCCCAAGATGTTGCTTCTCTTCAAGCTGTTTCT 311
Db 104 a|V|a|l|l|e|V|a|l|Y|A|s|P|I|n|U|a|r|g|S|e|r|P|r|o|A|G|I|n|U|S|e|r|L|e|U|A|r|g|I|n|U|A|S|e|r|H|V|A 124
OY 312 CACGTACTCTCGGTAATGAGAGAGCTTC-----AACTGTGTCACCTGCTTC 365
Db 312 C|A|C|G|T|A|C|T|C|T|G|G|T|A|A|T|G|A|A|G|A|G|A|G|C|T|T|C|---A|A|C|T|G|T|T|C|A|C|C|T|G|C|T|G|C 365

```

```

Db 124 l|S|e|r|L|e|u|V|a|l|G|I|n|A|l|e|u|A|r|g|A|S|n|A|l|a|G|I|n|U|a|r|g|H|r|a|S|P|I|e|C|y|S|e|L|e|U| 144
OY 366 AGGTGGTTTGTGTAAGTTCTCTGTTGTTTCCCTGCGCTGCTGAGAGAAATCACTT 425
Db 144 s|G|Y|G|Y|T|Y|G|I|n|U|a|r|g|P|H|e|S|e|r|S|e|r|G|I|n|U|Y|P|r|o|D|U|P|H|e|C|y|S|e|r|Y|S|H|r|Y|S|A|L|e 164
OY 426 A-----GTCCCACTCGATTTTCAGCT-----TGCTT 455
Db 164 u|A|A|l|a|l|e|P|r|o|P|r|o|V|a|l|P|r|o|P|r|o|S|e|r|A|l|A|H|G|I|n|U|P|r|o|L|e|u|A|S|P|L|e|u|A|S|P|C|y|S|e 184
OY 456 ACCTGTGTC-----AACATTCGCGCAACCCGAATTCCTCCCAATCTTTA 500
Db 184 r|S|e|r|C|y|S|e|L|Y|T|H|P|r|o|L|e|u|H|I|A|S|P|G|I|n|G|I|n|U|Y|P|r|o|V|a|l|G|I|n|U|I|L|e|u|P|r|o|H|e|U|Y 204
OY 501 TCTTGCTGCCAGCAGATGTCCTCAACAGAGCTGATACAGACAAATGGATGGTTA 560
Db 204 r|L|e|U|G|Y|S|e|r|A|Y|H|H|S|A|l|A|A|r|g|A|r|g|A|S|P|L|e|U|A|P|A|l|e|U|G|Y|L|e|T|H|A|l 224
OY 561 TGTGTTAAATGCCAGCTATACCTGTCCAAAGCTGATTCCTCCGAGTCAATTCCT 620
Db 224 A-----G|Y|G|Y|G|Y|G|I|n|U|a|r|g|P|H|e|S|e|r|S|e|r|G|I|n|U|Y|P|r|o|D|U|P|H|e|C|y|S|e|r|Y|S|H|r|Y|S|A|L|e 164
OY 621 GCGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTGCCGTGTTGCAAAATCAGT 680
Db 243 s|C|Y|S|I|L|e|P|r|o|V|a|l|G|I|n|U|a|S|P|H|I|S|Y|A|l|A|S|P|I|L|e|S|e|r|S|e|r|T|P|H|e|M|e|T|G|I|n|A|l|l 263
OY 681 AGATTCATTCAGAAAAGCAAAAGCTCCAAATGATGTGTCTAGTCAGCTGTTAGCTGG 740
Db 263 e|G|U|T|Y|T|L|e|P|r|A|V|a|l|Y|A|S|P|C|Y|A|r|G|I|Y|A|r|G|V|a|l|L|e|U|V|a|H|I|C|Y|S|G|I|n|A|G|I 283
OY 741 GATCTCCCTCGCCACCATGCTATGCTCAATCATATGAAGAGATGACATGCTTT 800
Db 283 Y|L|S|e|r|A|r|g|S|e|r|A|T|H|T|I|e|C|Y|S|e|U|A|T|Y|L|e|U|M|e|T|L|Y|S|e|r|V|a|l|A|r|g|L|e 303
OY 801 AGATGAGCTTACATTTGTGAAAAGAAAAGCACTATATATCTCAAACTTCAATT 860
Db 303 u|G|I|n|U|A|l|P|H|e|G|I|n|P|H|e|U|V|a|l|Y|S|G|I|n|A|r|g|S|e|r|I|L|e|S|e|r|P|r|o|A|S|P|H|e|S|e|r|H 323
OY 861 TCTGGCCAACTCTGTGATATGAGAAGATTTAAGAACAGATCGACATCGAGGCC 920
Db 323 e|W|e|r|G|Y|G|I|n|U|e|U|G|I|n|P|H|e|G|I|n|U|S|e|r|G|I|n|V|a|l|A|T|H|S|e|r|C|Y|S|A|l|A|l|G|U|A|l 343
OY 921 AAAGCAAACTCAAGCTGTCGACCTCGAAGAACCAAAATGAACCTGTCCCTGTGCTC 980
Db 343 a|A|L|S|e|r|-----P|r|o|S|e|r|G|Y|P|r|o|L|e|U|----- 350
OY 981 AGAGGTGACAGAAAAGCGAGCGCCCTGATCCAGCTGTGGCAGCTGTGTAATTC 1040
Db 351 -----GlyGluArgGlyLeuThrProAlaThrPro-----ThrSe 362
OY 1041 AGAGGACAGAGCAAAAGCCCGTCATCCGCCAGCGTCCCGCAGCTGCA 1100
Db 362 rGlnPheValIpheserPheProVal---SerValGlyValHisSerAlaProSerSerle 381
OY 1101 GCCGTGCTGTAGAGAGACGCCCGCTGTACAGCGCTCAGT 1143
Db 381 uProTyLeu-----HisSerProIleThrThrSerProSer 393

```

RESULT 9

```

US-09-614-124B-805
/ Sequence 805, Application US/09614124B
/ Patent No. 6630574
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darick
/ APPLICANT: Recter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ DIAGNOSIS OF LUNG CANCER

```

```
FILE REFERENCE: 210121.478C9
CURRENT APPLICATION NUMBER: US/09/614,124B
CURRENT FILING DATE: 2001-07-11
NUMBER OF SEQ ID NOS: 1668
SOFTWARE: FastSeq for Windows Version 3.0.
SEQ ID NO 805
LENGTH: 394
TYPE: PRT
ORGANISM: Homo sapiens
US-09-614-124B-805

Alignment Scores:
Pred. No.: 1,25e-32 Length: 394
Score: 455.00 Matches: 122
Percent Similarity: 49.87% Conservative: 75
Best Local Similarity: 30.89% Mismatches: 143
Query Match: 12.59% Indels: 55
DB: Gaps: 13

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-614-124B-805 (1-394)

QY 45 GGTGGCTGCTGCGGAAAGTGAACGAA-----AAAGT 77
DB 26 GlyGlyAlaGlyGly-SerGlySerHisGlyThrLeuGlyLeuProSerGlyGlySerCy 45
QY 78 GGTGTAATTTGAATGACCGGCGCATTTGTGAATACATCCACATTTTGAAGCCAT 137
DB 45 GLeuLeuLeuSerCyArgProPheLeuAlaHisSerAlaGlyTyrTleLeuGlySerVa 65
QY 138 TAATATCAATGCTGCTCAAGCTTATGAACGAGAGTTGCAACAGACAGACAAAGTTTAATTAC 197
DB 65 IAsnValArgCyAsnThrTleValArgArg---AlaGlyGlySerValSerLeuG 84
QY 198 AGAGCTTCATC-----CAGCATTCAGCGAAACATTAAGTTGACATTGATTCAGTCAGAA 251
DB 84 uGlnIleLeuProAlaGlyGlyGlyValAlaArgAlaArgLeuArgSerGlyLeuTyrSerAl 104
QY 252 GGTTTAGTTAGTTAGTGAAGCTCCCAAGATGTTGCTCTCTCTTCAGACTGTTTCT 311
DB 104 aValIleValTyrAspIleArgSerProArgAlaGlySerLeuArgGlyAspSerThrVa 124
QY 312 CACTGTACTCTGGTGAACCTGGAAGAGCTTC-----AAGCTGTTCACTGCTTGC 365
DB 124 IserLeuValValGlnAlaLeuArgArgAsnAlaGlyArgThrAspIleCySerLeuLeu 144
QY 366 AGGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATTCACCTCT 425
DB 144 sGlyGlyTyrGlyArgPheSerSerGlyTyrProGluPheCySerTyrThrTyrAlaLe 164
QY 426 A-----GTCCCTACCTGCAATTTCTCAGCTT-----TGCTT 455
DB 164 uAlaAlaIleProProProValProProSerAlaThrGluProLeuAspLeuAspCySe 184
QY 456 ACCTGTGTC-----AACATTGGGCGCAACCGGAATTTCTTCCCATCTTAA 500
DB 184 rSerCySerGlyThrProLeuHisArgGlnGlyProValGlnIleLeuProLeuTyr 204
QY 501 TCTTGCTGCGCAGAGATGCTCTCAACAGAGCTGATACAGACAGAAATGGATTGTTA 560
DB 204 rLeuGlySerAlaTyrHisAlaAlaArgArgAspMetLeuAspAlaLeuGlyTleThrAl 224
QY 561 TGTGTTAAATCCAGCTATACCTGTCCAAAGCTGACTTTATCCCGAGCTCATTTTCT 620
DB 224 aLeuLeuAsnValSerSerAspCyProAsn---HisPheGlnGlyHisTyrGlnTyrTyr 243
QY 621 GCGTGGCTGTGTAATGACAGCTTTTGTGAATAATTTCGCTGTTGGAACAATCAGT 680
DB 243 sCySileProValGlnAspAsnHisArgAlaAspIleSerSerTyrPheMetGluAla 263
QY 681 AGATTTCAATGAGAAAGCAAGCTTCGAATGATGTGTTCTAGTCACTGTTTAACTGG 740
DB 263 eGlnTyrTleAspAlaValTyrAspCyArgGlyArgValTleValHisCySerGlnAlaG 283
```

```
QY 741 GATCTCCCGCTCGCCACCATGCTATCCGCTACATCATGAGAGATGACATGCTTT 800
DB 283 yIleSerArgSerAlaThrTleCySerLeuAlaTyrLeuMetMetLeuValArgValArgLe 303
QY 801 AGATGAGCTTACAGATTGTGGAAGAAAGAAAGCACTATATCTTCAATTCAATT 860
DB 303 uGlnGlnAlaPheGlnPheValTyrGlnArgSerTleIleSerProAsnPheSerPh 323
QY 861 TCTGGCCCACTCTTGTGATATGAGAAGAAAGTTAAGAACAGACTGAGACATCAGGCC 920
DB 323 eMetGlyGlnLeuLeuGlnPheGlnSerGlnValAlaThrSerCyAlaAlaGluAl 343
QY 921 AAAGCAATATCAAGCTGCTGCACTGGAAGAGCAAAATGAACCTGCTGCTCTC 980
DB 343 aAlaSer-----ProSerGlyProLeu----- 350
QY 981 AGAGGTGACAGAAAGAGAGAGCGCCCTCAGTCACCTGTGCGGACTGTGTACTTC 1040
DB 351 -----GlyGlyArgGlyTyrThrProAlaThrPro-----ThrSe 362
QY 1041 AGAGGACAGACAGAAAGCGCTGATCCCGGACGCTGCCAGCTGCCAGCGTCA 1100
DB 362 rGlnPheValPheSerPheProVal---SerValGlyValHisSerAlaProSerSerle 381
QY 1101 GCGTGGCTGTTAGAGACAGCGCGCTGGTACAGGCGCTCAGT 1143
DB 381 uProTyrLeu-----HisSerProIleThrThrSerProSer 393

RESULT 10
US-09-671-325-805
Sequence 805, Application US/09671325
Patent No. 6667154
GENERAL INFORMATION:
APPLICANT: Wang, Tongfong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedrick, Tom
APPLICANT: Carter, Darriick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C12
CURRENT APPLICATION NUMBER: US/09/671,325
CURRENT FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 1825
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 805
LENGTH: 394
TYPE: PRT
ORGANISM: Homo sapiens
US-09-671-325-805

Alignment Scores:
Pred. No.: 1,25e-32 Length: 394
Score: 455.00 Matches: 122
Percent Similarity: 49.87% Conservative: 75
Best Local Similarity: 30.89% Mismatches: 143
Query Match: 12.59% Indels: 55
DB: Gaps: 13

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-671-325-805 (1-394)

QY 45 GGTGGCTGCTGCGGAAAGTGAACGAA-----AAAGT 77
DB 26 GlyGlyAlaGlyGly-SerGlySerHisGlyThrLeuGlyLeuProSerGlyGlySerCy 45
QY 78 GGTGTAATTTGAATGACCGGCGCATTTGTGAATACATCCACATTTTGAAGCCAT 137
DB 45 GLeuLeuLeuSerCyArgProPheLeuAlaHisSerAlaGlyTyrTleLeuGlySerVa 65
```



```

Db 204 rLeuGlySerAlaTyHAlaAlaArgArGAspMetLeuAspAlaLeuGlyIleThrAl 224
QY 561 TGTTGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTATCCCGAGTCTCATTTCTT 620
Db 224 aLeuLeuAspValSerSerAspCysProAsn---HisPheGluGlyHisTyGlnTyIle 243
QY 621 GCCTGTGCTGTGATGACAGCTTTTGTGAAATAATTTCCGTGTGTCGACCAATACAGT 660
Db 243 sCyIleProValGluAspAsnHisValAspIleSerSerTrpPheMetGluAlaIle 263
QY 681 AGATTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTGTTCTAGTGACATGTTAGCTGG 740
Db 263 eGlnTyIleAspAlaValIleValAspCysArgGlyArgValIleValHisCysGlnIle 283
QY 741 GATCTCCCGCTCCGCCCACTGCTATCGCTATCATATGAAAGATGACATGCTTTT 800
Db 283 yIleSerArgSerAlaThrIleCysLeuAlaTyLeuMetMetIleValArgValArgLe 303
QY 801 AGATGAAGCTTACAGATTGTGAAAGAAAGAAAGCCTATATCTCCAAATCTCAATTT 860
Db 303 uGlnGluAlaPheGluPheValIleValArgArgSerIleIleSerProAsnPheSerPn 323
QY 861 TCTGGCCAACTCCTGTGACTATGAGAAAGATTAAAGAACAGATGAGCATCAGGGCC 920
Db 323 eMetGlyGlnLeuLeuGlnPheGlnSerGlnValIleAlaThrSerCysAlaIleGluAl 343
QY 921 AAAGCAAACTCAAGCTGCTGCACTGAGAAAGCAATGAACCTGCTCCCTGCTGCTC 980
Db 343 aAlaSer-----ProSerGlyProLeu----- 350
QY 981 AGAGGGTGCAGAAAGCAAGAGCGCCCTCAGTCCACCTGTGCGAGCTCTGTAACCTC 1040
Db 351 -----GlyGluArgGlyIleThrProAlaThrPro-----ThrSe 362
QY 1041 AGAGGCAAGCAAGAGCGCCCTGTCATCCCGCAGCGTCCGAGCGTCCGACGCTGCA 1100
Db 362 rGlnPheValPheSerPheProVal---SerValGlyValHisSerIleProSerSerIle 381
QY 1101 GCCGTGCTGTAGAGACAGCGCCGCTGTGTAACAGCGCTCACT 1143
Db 381 uProTyIleu-----HisSerProIleThrThrSerProSer 393

```

RESULT 12

```

US-08-530-290-23
/ Sequence 23, Application US/08530290
/ Patent No. 5958721
/ GENERAL INFORMATION:
/ APPLICANT: Marshhall, Christopher John
/ APPLICANT: Ashworth, Alan
/ APPLICANT: Hughes, David Anthony
/ TITLE OF INVENTION: Methods for Screening of Substances for
/ TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
/ NUMBER OF SEQUENCES: 24
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSER: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/530,290
/ FILING DATE: 14-DEC-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO PCT/GB94/00694
/ FILING DATE: 31-MAR-1994
/ PRIOR APPLICATION DATA:

```

```

/ APPLICATION NUMBER: GB 9402573.1
/ FILING DATE: 10-FEB-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9307250.2
/ FILING DATE: 07-APR-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Baetian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 084611-0000000US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 23:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 394 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-530-290-23

```

Alignment Scores:

	Pred. No.:	6,62e-32	Length:	394
Score:	447.00	Matches:	121	
Percent Similarity:	49.62%	Conservative:	75	
Best Local Similarity:	30.63%	Mismatches:	144	
Query Match:	12.37%	Indels:	55	
DB:	2	Gaps:	13	

```

US-10-029-345a-108_copy_538_2532 (1-1995) x US-08-530-290-23 (1-394)
QY 45 GGTGGCTGCTGCTGGAAGTGGAAAGCA-----AAAGT 77
Db 26 GlyGlyAlaGlyGly-SerGlySerHisGlyThrLeuGlyLeuProSerGlyGlyIlePcy 45
QY 78 GCTGCTAATGATAGAGCGCCCAATTTGTGAATACATACATCCACATTTTGGAGCCAT 137
Db 45 sLeuLeuLeuAspCysArgProPheLeuAlaHisSerAlaGlyTyIleLeuGlySerVal 65
QY 138 TAATATCAACTGCTCCAAAGCTTATGAGCAAGAGTTCGACAGACAGACAAAGTTAATTAC 197
Db 65 lAnValArgCysAsnThrIleValArgArgArg---AlaTyGlySerValSerLeuG 84
QY 198 AGAGCTCATC-----CAGCATTCAGCGAAACTAAGCTTGACATGATTGACGTGAGAA 251
Db 84 uGlnIleuProAlaGluGluGluValArgAlaArgLeuArgSerGlyLeuTySerAl 104
QY 252 GGTGTAGTTTACGATCAAGCTCCCAAGATGTGCTCTCTTTCAGACTGTTTCT 311
Db 104 aValIleValIlePheGluArgSerProArgAlaGluSerIleuArgGluAspSerThrVal 124
QY 312 CACTGACTTCTGGGTAACTGAGAAAGACTTC-----AACTGTTCACCTGCTGTC 365
Db 124 lSerLeuValValGlnAlaLeuArgArgAsnAlaGluArgThrAspIleCysLeuLeu 144
QY 366 AGTGGGTTTGTGAGATTCTCTGTTGTTCCCTGCGCTCTGGAAGAAATCACTCT 425
Db 144 sGlyGlyTyGluArgPheSerSerGluTyProGluPheCysSerLeuThrValAla 164
QY 426 A-----GTCCCTACTGCACTTCTCAGACT-----TGCTT 455
Db 164 uAlaAlaIleProProProValProProSerAlaThrGluProLeuAspLeuAspCysSe 184
QY 456 ACCTGTGTC-----AACATGGGCAACCCGAATTTCTCCAAATCTTTA 500
Db 184 rSerCysGlyThrProLeuHisAspGlnGluProValGlnIleLeuProPheLeu 204
QY 501 TCTTGGCTGCGCAGAGATGCTCTCAACAAGAGCTGATACAGCAGAAATGGGATTGTTA 560
Db 204 rLeuGlySerAlaTyHAlaAlaArgArgAspMetLeuAspAlaLeuGlyIleThrAl 224
QY 561 TGTTGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTATCCCGAGTCTCATTTCT 620

```


Db 224 aleuLeuAenValSerSerAspCysProAsn---HisPheGluGlyHisTyrglnTyly 243
QY 621 GCCTGTGCTGCTGAGATGACGCTTTTGTGAGAAATTTTGCCTGTGTGACAAATCAGT 680
Db 243 sCysIleProValGluAspAsnHisLysAlaAspIleSerSerTrpPheMetGluAlaI 263
QY 681 AGATTTCATTGAGAAAGCAAAAGCCCTCAATGATGATGTGTTCTAGTGCATGTTTACTG 740
Db 263 eGluTyrlleAspAlaValLysAspCysArgGlyArgValLeuValHisCysGlnAlaG 283
QY 741 GATCTCCGCTCCGCAACATGCTATGCTTACATCATGATGAGAGATGAGATGATGCTTT 800
Db 283 yIleSerArgSerIleAlaTrilleCysLeuAlaTyrlleuMetLysLysArgValArgLe 303
QY 801 AGATGAAGCTTACAGATTGTGAAAGAAAGAAAGCACTATATCTCCAACTTCAATT 860
Db 303 uGlnGluAlaPheGluPheValLysGlnArgAspSerIleIleSerProAsnPheSerPh 323
QY 861 TCTGGGCCAACTCCTGCACTATGAGAAAGATTAAGAACAGACTGTGAGATCAGGGCC 920
Db 323 eMetGlyGlnLeuLeuGlnPheGlnSerGlnValIleuAlaThrSerCysAlaAlaGluAl 343
QY 921 AAAGAGCAAACTCAAGCTGCTGCACTGTGAGAAAGCAATGAACTGCTGCTGCTC 980
Db 343 aAlaSer-----ProSerGlyProLeu----- 350
QY 981 AGAGGCTGACAGAAAGAGAGAGCCCTCACTGATCCACCTGTGCGCATCTGTACTCT 1040
Db 351 -----ArgGluArgGlyLysTrpProAlaThrPro-----Thre 362
QY 1041 AGAGGAGAGAGAGCAAAAGCCGCTGCATCCCGCAGCGTGCGCGAGCGTGCAGCGTGA 1100
Db 362 rGlnPheValPheSerPheProVal---SerValGlyValHisSerAlaProSerSerle 381
QY 1101 GCCGTGCTGTGTAAGAGACGCCGCTGTGATCAGAGCCCTCAGT 1143
Db 381 uProTyrlleu-----HisSerProIleThrTrpSerProSer 393

RESULT 13
US-09-702-705-827
; Sequence 827 Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedrick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: DIAGNOSIS OF LUNG CANCER
; CURRENT APPLICATION NUMBER: US/09/702,705
; NUMBER OF SEQ ID NOS: 2000-10-30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 827
; LENGTH: 394
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-702-705-827

Alignment Scores:
Pred. No.: 6.62e-32 Length: 394
Score: 447.00 Matches: 122
Percent Similarity: 50.38% Conservative: 77
Best Local Similarity: 30.89% Mismatches: 141
Query Match: 12.37% Indels: 55
DB: 4 Gaps: 14

US-10-029-345a-108_copy_538_2532 (1-1995) x US-09-702-705-827 (1-394)
QY 45 GGTGGCTGCTGCTGAGAAAGCAACGAA-----AAAGT 77
Db 26 GLyAlaGlyGly---SerGlySerHisGlyThrLeuGlyLeuProSerGlyGlyGlyScy 45
QY 78 GCTGTAAATTGATTAACCGGCAATTTGTGATTAACATATACATCCCACTTTTGAACCAT 137
Db 45 sLeuLeuLeuAspCysArgProPheLeuAlaHisSerAlaGlyTyrlleuGlySerVa 65
QY 138 TAATATCAACGTCTCCAAAGCTTATGAAAGCAAGAGTTGCAACAGCAAAAGTTTATAC 197
Db 65 lAsnValArgCysSerThrIleValAlaArgArgArg---AlaLysGlySerValSerLeuG 84
QY 198 AGAGCTATC-----CAGCATTCAGCAAAACATAAGTTGACATTGATTGACAGACAGA 251
Db 84 uGlnIleLeuProAlaGlnGluGluValAlaArgAlaArgLeuArgSerGlyLeuTySerAl 104
QY 252 GATTGTAGTTTAAAGATGAAGCTCCCAAGATGTTGCTCTCTCTTCACTGTTTCT 311
Db 104 aValIleValIleValIlePheIleuArgSerProAlaArgAlaGlnSerLeuAlaGlnAspSerThrVa 124
QY 312 CACTGTACTTCTGGGTAAACTGTGAGAAAGACTTC-----AACTGTCTCACTGCTTGC 365
Db 124 lSerLeuValValGlnAlaAlaLeuArgArgAsnAlaGlnArgThrAspIleCysLeuLeu 144
QY 366 AGGTGGGTTTGTGAGTTCTCTGTTTCTGCTGCTGCTGCTGCTGAGAAAGTCACTCT 425
Db 144 sGlyGlyTyrgluAlaArgPheSerSerGlyTyrglnPheCysSerLysThrLysAlaLe 164
QY 426 A-----GTCCCTTACCTGCATTTCTCAGCCT----- 450
Db 164 uAlaAlaIleProProProValProProSerAlaThrGlnProLeuAspLeuGlyCysSe 184
QY 451 ----TGC---TTACTGTGTGCAACATT---GGGCCAACCCGAATTTTCCCAATCTTA 500
Db 184 rSerCysGlyThrProLeuHisAspGlnGlyGlyProValGlnIleLeuProPheLeuTy 204
QY 501 TCTTGGCTGCCAGAGATGCTTCTCAACAGAGAGCTATACAGCAGAAATGATGCTTA 560
Db 204 rLeuGlySerAlaTyrlleAlaAlaArgArgPheMetLeuAspAlaLeuGlyIleThrAl 224
QY 561 TGTGTAAATGCCAGATATACCTGTCCAAAGCTGATTTATCCCGAGTCAATTTCT 620
Db 224 aleuLeuAsnValSerSerAspCysProAsn---HisPheGluGlyHisTyrglnTyly 243
QY 621 GCCTGTGCTGCTGAGATGACGCTTTTGTGAGAAATTTTGCCTGTGTGACAAATCAGT 680
Db 243 sCysIleProValGluAspAsnHisLysAlaAspIleSerSerTrpPheMetGluAlaI 263
QY 681 AGATTTCATTGAGAAAGCAAAAGCCCTCAATGATGATGTGTTCTAGTGCATGTTTACTG 740
Db 263 eGluTyrlleAspAlaValLysAspCysArgGlyArgValLeuValHisCysGlnAlaG 283
QY 741 GATCTCCGCTCCGCAACATGCTATGCTTACATCATGATGAGAGATGAGATGATGCTTT 800
Db 283 yIleSerArgSerIleAlaTrilleCysLeuAlaTyrlleuMetLysLysArgValArgLe 303
QY 801 AGATGAAGCTTACAGATTGTGAAAGAAAGAAAGCACTATATCTCCAACTTCAATT 860
Db 303 uGlnGluAlaPheGluPheValLysGlnArgAspSerIleIleSerProAsnPheSerPh 323
QY 861 TCTGGGCCAACTCCTGCACTATGAGAAAGATTAAGAACAGACTGTGAGATCAGGGCC 920
Db 323 eMetGlyGlnLeuLeuGlnPheGlnSerGlnValIleuAlaThrSerCysAlaAlaGluAl 343
QY 921 AAAGAGCAAACTCAAGCTGCTGCACTGTGAGAAAGCAATGAACTGCTGCTGCTC 980
Db 343 aAlaSer-----ProSerGlyProLeu----- 350
QY 981 AGAGGCTGACAGAAAGAGAGAGCCCTCACTGATCCACCTGTGCGCATCTGTACTCT 1040
Db 351 -----ArgGluArgGlyLysTrpProAlaThrPro-----Thre 362


```

; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 827
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-614-124B-827

Alignment Scores:
Pred. No.: 6,62e-32      Length: 394
Score: 447.00           Matches: 122
Percent Similarity: 50.38% Conservative: 77
Best Local Similarity: 30.89% Mismatches: 141
Query Match: 12.37%      Indels: 55
DB: 4                   Gaps: 14

US-10-029-345A-108_COPY_538_2532 (1-1995) x US-09-614-124B-827 (1-394)

QY 45 GGTGGCTCTGCTGGAAGAGGAAACGGA-----AAAGT 77
DB 26 GlyGlyAlaGlyGly-SerGlySerHisGlyThrLeuGlyLeuProSerGlyGlyLeuCy 45
QY 78 GGTGCTAATTTAGACCGGCATTTGTGGAATACATACATCCACATTTTGGAGCCAT 137
DB 45 sLeuLeuLeuAspCyHisGProPheLeuLHisSerAlaGlyTyrLLeuGlySerVa 65
QY 138 TAAATCACTGCTCCAGCTTATGAGCGAAGTTGCAACGAAAGTTTAAATAC 197
DB 65 LAsnValArgCyAsnThrLLeValArgArgArg--AlaLysGlySerValSerLeuG 84
QY 198 AAGAGCTCATC-----CAGCATTCAGCGAAACATTAAGTTGACATTGTCAGTCAGAA 251
DB 84 uGlnIleLeuProAlaGluGluGluValArgAlaArgLeuArgSerGlyLeuTyrSerAl 104
QY 252 GGTGTAGTTAGTACATCAAGCTCCCAAGATGTGCTCTCTCTCTGACACTGTTTCT 311
DB 104 aValLLeValTyrzAspGluArgSerProArgAlaGluSerLeuArgGluAspSerThrVa 124
QY 312 CACTGTACTTCTGGGTAAACTGGAGAAGACTTC-----AACTGTCTCACTGCTTGC 365
DB 124 LserLeuValValGlnAlaLeuArgAsnAlaGluArgThrAspLLeCysLeuLeuL 144
QY 366 AGGTGGTTTGCTAGTCTCGTTCGTTTCCCTGGCTGCTGTAAGAAATCCACTCT 425
DB 144 sGlyGlyTyrGluArgPheSerSerGlyTyrProGluPheCysSerLysThrLysAlaLe 164
QY 426 A-----GTCCCTACCTGCATTTCTCAGCT----- 450
DB 164 uAlaAlaIleProProProValProProSerAlaThrGluProLeuAspLeuGlyCysSe 184
QY 451 -----TGC---TTACCTGTGGCCACATC--GGGCCAACCCGAATTTTCCCAATCTTTA 500
DB 184 LserCySglYThrProLeuHisAspGlnGlyGlyProValGluIleLeuProPheLeuT 204
QY 501 TCTTGGCTGCACGAGATGCTCTCAACAGAGGCTGATACAGACGAATGGGATTTGTTA 560
DB 204 rLeuGlySerAlaTyrHisAlaAlaArgArgAspMetLeuAspAlaLeuGlyLLeThrAl 224
QY 561 TGTGTTAAATGCACCTATACCTGTCTCAAGCTGACTTATCCCGAGTCTCATTTTCT 620
DB 224 aLeuLeuAsnValLserSerAspCySProAsn--HisPheGluGlyHisTyrGlnTyrL 243
QY 621 GCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTGGCGGTGTTGACAAATCAGT 680
DB 243 sCysLLeProValGluAspAsnHisLysAlaAspLLeSerSerTrpPheMetGlnAlaI 263
QY 681 AGATTCTTGAAGAAAGAAAGCCTCAATGATGATGTTTATAGGACCTGTTAGCTGG 740
DB 263 eGlnTyrLLeAspAlaValLysAspCySArgGlyArgValLeuValHisCySglnAlaG 283
QY 741 GATCTCCGCTCCGACCATCGCTATGCTCATCATGAGAGATGAGCATGTCTTT 800
DB 283 yLlSerArgSerAlaThrLLeCysLeuAlaTyrLeuMetMetLysLysArgValArgLe 303

```

```

QY 801 AGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCCTACTATATCTCAACTTCATTT 860
DB 303 uGlnGluAlaPheGluPheValLysGlnArgArgSerLLeSerProAsnPheSerP 323
QY 861 TCTGGCCCAACTCTGTGACATATGAGAAAGATTAAGAACCACTGACCATCAGGGCC 920
DB 323 eMetGlyGlnLeuLeuGlnPheGlnSerGlnValLLeuAlaThrSerCySAlaAlaGluAl 343
QY 921 AAAGAGCAACTCAGCTGTGCACTCGAGAGAGCCAAATGAACCTGCTCCCTGCTCTC 980
DB 343 aAlaSer-----ProSerGlyProLeu-----ThrSe 362
QY 981 AGAGGTGACAGAAAGACGAGACGCCCTGACGTCAACCTGTGCCACTGTCTACCTC 1040
DB 351 -----ArgGluArgGlyLysThrProAlaThrPro----- 362
QY 1041 AGAGGACGACGACAAAGCCCGTGCATCCCGCAGCGTCCAGCGTGCACGCTGCA 1100
DB 362 rGlnPheValPheSerPheProVal--SerValGlyValHisSerAlaProSerSerLe 381
QY 1101 GCCGTGCTGTATAGAGGACAGCCCGCTGTATACAGGGCGCTCAGT 1143
DB 381 uProTyrLeu-----HisSerProIleThrThrSerProSer 393

```

Search completed: June 21, 2004, 12:43:22
 Job time : 42.4447 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 12:17:10 ; Search time 12558 Seconds

(without alignments)
6885.611 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccaatgagatgattg9.....tggaatcatgaggtctcc 1995Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_or:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrc:*
38: em_sy:*
39: em_hugo_hum:*
40: em_hugo_mus:*
41: em_hugo_other:*
```

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1995	100.0	5450	6	AX482439
2	1995	100.0	5450	6	AX482478
3	1991.8	99.8	1998	6	AX260342
4	1991.8	99.8	3059	6	AX278461
5	1991.8	99.8	3496	6	AX441210
6	1991.8	99.8	3521	9	AB052156
7	1991.8	99.8	3544	6	AX260340
8	1991.8	99.8	3566	9	AF506796
9	1991.8	99.8	3766	6	AX374994
10	1991.8	99.8	4790	6	BD171157
11	1991.8	99.8	4790	6	BD183422
12	1991.8	99.8	4790	9	AB051487
13	1990.2	99.8	2732	6	AX180875
14	1990.2	99.8	3104	6	AX405700
15	1988.6	99.7	2102	6	AX713989
16	1988.6	99.7	2102	9	AX055973
17	1988.6	99.7	5111	6	AX482372
18	1968.8	98.7	2118	6	AX099933
19	1963.8	98.4	3284	6	BC042101
20	1962.2	98.4	2071	6	AX921917
21	1906.8	95.6	2807	6	AX921915
22	1844.2	92.4	2807	9	BC031643
23	1653.8	82.9	1935	9	AY038927
24	1653.8	82.9	3332	6	AX441229
25	1572.4	78.8	4943	10	BC059232
26	1572.4	78.8	4975	10	BC057321
27	1567.6	78.6	4874	10	AB052157
28	1554.8	77.9	2756	6	AX482444
29	1531	76.7	4992	10	AF345951
30	1200.6	60.2	4827	10	AF345952
31	1183.4	59.3	172206	9	AC007619
32	1183.4	59.3	188344	2	AC131617
33	879.4	44.1	1916	6	AX835305
34	879.4	44.1	1916	6	AX098310
35	858.4	43.0	201474	2	AC126692
36	858.4	43.0	236589	2	AC118035
37	848.6	42.5	242590	2	AC133722
38	848.6	42.5	244605	2	AC097818
39	848.6	42.5	250782	2	AC128093
40	848.6	42.5	250782	2	AC128093
41	727.6	36.5	749	6	AX868469
42	714.2	35.8	4026	10	AF345953
43	405.8	20.3	3221	5	BC063941
44	387.8	19.4	2377	6	AR255944
45	387.8	19.4	2377	9	HS027193

ALIGNMENTS

```
RESULT 1
AX482439          5450 bp   DNA
DEFINITION       Sequence 108 from Patent WO02057460.
ACCESSION        AX482439
VERSION          AX482439.1  GI:22316984
KEYWORDS
SOURCE
ORGANISM         Homo sapiens (human)
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE        1 Todderud,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G.,
                  Suchard,S., Banas,D., Baasolino,D., Feder,J., Krystek,S.,
                  Meatee,P., Mintier,G., Siemers,N., Jackson,D.G. and Ramanathan,C.
```

TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 108 25-JUN-2002;
FEATURES SQUIBB BRISTOL MYERS CO (US)
source Location/Qualifiers

1..5450
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
538..2535
/note="unamed protein product"
/codon_start=1
/protein_id="CAD44451.1"
/db_xref="GI:22316985"
/translation="NAHEMIGQIVTERIVALLSETEKVLIDSRPVEYNTSHILE
FLTVLIGLKEKSNVHLIAGPABFRCFPGCEKGTLPVCSOCPLEVANIGT
RLPNTLYGCGQKDVNKEIQQNGIGYVLMASYCPKPDFIESHPLRVPNDSRCK
KRPDISPNFLEKAKASNCVLAHCLAGISRSATIAIAYIMKMDSLDRAHYVE
TNLPSPCADSATSEAGRPVHPASVPSVOPSLDEDFVLVALGSLHLSADRLD
SNLKRFSFLDIKSYVSASMAAHLHGSSSEDLRYKESITLDGNNLQPSVOE
LSBOIETSPDKKEASIPKLOTPAPSPDSQSKRLHVSRTSSGTGSLSPHRSQS
VEDNHTSFLRGLSTSOCHLTKSAGLGMHSDILAPORTSBLTSSMTFATESHP
YSASAIYGSASYSAYSCSQPTCDQYVSVRRQKSDRADSRSWHEBSPREKQFK
RRSCMERBESTMSNRSEBELKRVGSQSPSGSMETIEVS"

ORIGIN

Query Match 100.0%; Score 1995; DB 6; Length 5450;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGCCCATGAGATGATGGAACCTCAATTGTTACTGAGAGTGTGCTCTGTGAA 60
Db 538 ATGCCCATGAGATGATGGAACCTCAATTGTTACTGAGAGTGTGCTCTGTGAA 60
QY 61 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCGCGCATTTGTGAATCAATACATCC 120
Db 598 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCGCGCATTTGTGAATCAATACATCC 120
QY 121 CACATTTGGAAGCATTAATATCACTGCTCCAACTTATGAGGAAAGTTGCAACG 180
Db 658 CACATTTGGAAGCATTAATATCACTGCTCCAACTTATGAGGAAAGTTGCAACG 180
QY 181 GACAAAGTGAATTAACAGAGCTCATCCAGCAATGAGGAAACATTAAGGTTGACATGAT 240
Db 718 GACAAAGTGAATTAACAGAGCTCATCCAGCAATGAGGAAACATTAAGGTTGACATGAT 240
QY 241 TGCAGTCAGAAAGGTGATGATTCAGATCAAGACTCCAAAGTGTGCTCTCTCTTCA 300
Db 778 TGCAGTCAGAAAGGTGATGATTCAGATCAAGACTCCAAAGTGTGCTCTCTCTTCA 300
QY 301 GACTGTTTTCTCACTGTAATCTTGGGTAACCTGAGAGAGACTTCACTGTTCACCTG 360
Db 838 GACTGTTTTCTCACTGTAATCTTGGGTAACCTGAGAGAGACTTCACTGTTCACCTG 360
QY 361 CTTCGAGGAGGTTGCTGAGATCTCTGTGTTTCCCTGAGCTCTGTGAGAGAAATCC 420
Db 898 CTTCGAGGAGGTTGCTGAGATCTCTGTGTTTCCCTGAGCTCTGTGAGAGAAATCC 420
QY 421 ACTCTAGTCCCTACCTGCAATTCCTGAGCTGCTTACCTGTTGCAACATTTGGGCAAC 480
Db 958 ACTCTAGTCCCTACCTGCAATTCCTGAGCTGCTTACCTGTTGCAACATTTGGGCAAC 480
QY 481 CGAATTCCTCCCATCTTATCTTGGCTGCTGAGAGATGCTCTCAACAAGAGAGCTGATA 540
Db 1018 CGAATTCCTCCCATCTTATCTTGGCTGCTGAGAGATGCTCTCAACAAGAGAGCTGATA 540
QY 541 CAGCAGAAATGGATGTTATGTTAAATGCAAGCTATACCTGTCCAAAGCTGACTTT 600
Db 1078 CAGCAGAAATGGATGTTATGTTAAATGCAAGCTATACCTGTCCAAAGCTGACTTT 600
QY 601 ATCCCGAGTCAATTCCTGCTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 660
```

```

Db 1138 ATCCCGAGTCAATTCCTGCTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 1197
QY 661 CCGAGTGTGGAACAATCAGTAATTTCAATGAGAAAGCAAGCTCCATGATGATGTT 720
Db 1198 CCGAGTGTGGAACAATCAGTAATTTCAATGAGAAAGCAAGCTCCATGATGATGTT 720
QY 721 CTAGTGCACTGTTAGCTGGGATCTCCGCTCCGCAACATCCGTATCCCTATCATG 780
Db 1258 CTAGTGCACTGTTAGCTGGGATCTCCGCTCCGCAACATCCGTATCCCTATCATG 780
QY 781 AAGAGATGGAACATGCTCTTGAATGAGCTTACAGATTTGTGAAAGAAAGACCTACT 840
Db 1318 AAGAGATGGAACATGCTCTTGAATGAGCTTACAGATTTGTGAAAGAAAGACCTACT 840
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGACATATGAGAAAGATTAAGAAC 900
Db 1378 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGACATATGAGAAAGATTAAGAAC 900
QY 901 CAGACTGAGCATAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTGAGAGAGCAAAAT 960
Db 1438 CAGACTGAGCATAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTGAGAGAGCAAAAT 960
QY 961 GAACCTGTCCCTGCTGCTCAGAGGGTGAAGAAAGGAGACCGCTCTCAGTCCACC 1020
Db 1498 GAACCTGTCCCTGCTGCTCAGAGGGTGAAGAAAGGAGACCGCTCTCAGTCCACC 1020
QY 1021 TGTGCGCATCTGTCTATCTCAGAGGCAAGAGCAAAAGCCCGTGCATCCGCAAGCTG 1080
Db 1558 TGTGCGCATCTGTCTATCTCAGAGGCAAGAGCAAAAGCCCGTGCATCCGCAAGCTG 1080
QY 1081 CCAAGGCTGCCAGCTGAGCGCTGCTGCTGATGAGAGCAACCCGCTGTGACAGGCTC 1140
Db 1618 CCAAGGCTGCCAGCTGAGCGCTGCTGCTGATGAGAGCAACCCGCTGTGACAGGCTC 1140
QY 1141 AGTGGGCTGCACTGTCCGCAAGCAGTGAAGACAGCAATTAAGCTCAAGCTTCTCTC 1200
Db 1678 AGTGGGCTGCACTGTCCGCAAGCAGTGAAGACAGCAATTAAGCTCAAGCTTCTCTC 1200
QY 1201 TCTTGATATCAATTCAGATTCATATTCAGCCAGCATGGAGATCTTACATGCTTTC 1260
Db 1738 TCTTGATATCAATTCAGATTCATATTCAGCCAGCATGGAGATCTTACATGCTTTC 1260
QY 1261 TCCCTATCAAGATGCTTTGAAATTAATAAACTTCACTACTCTGATGAGGACCAAC 1320
Db 1798 TCCCTATCAAGATGCTTTGAAATTAATAAACTTCACTACTCTGATGAGGACCAAC 1320
QY 1321 AAGCTATGCCAGTCTCCCTGTTGAGGAATTCAGAGAGCTGCGGAAACCAAGTCT 1380
Db 1858 AAGCTATGCCAGTCTCCCTGTTGAGGAATTCAGAGAGCTGCGGAAACCAAGTCT 1380
QY 1381 GATAAGAGAGAGCAGCATCCCAAGAAAGCTGCAAGCCGAGGCTTCAAGAGCCAG 1440
Db 1918 GATAAGAGAGAGCAGCATCCCAAGAAAGCTGCAAGCCGAGGCTTCAAGAGCCAG 1440
QY 1441 AGCAAGCATTTGATTCGATCAGAAACAGACAGATGAGGACCGCCAGAGCTCTTTTA 1500
Db 1978 AGCAAGCATTTGATTCGATCAGAAACAGACAGATGAGGACCGCCAGAGCTCTTTTA 1500
QY 1501 TCTCCACTGATGAAAGGAGGCTGAGAGCAATTAACAACAGCTTCTTTTGGGC 1560
Db 2038 TCTCCACTGATGAAAGGAGGCTGAGAGCAATTAACAACAGCTTCTTTTGGGC 1560
QY 2038 TCTCCACTGATGAAAGGAGGCTGAGAGCAATTAACAACAGCTTCTTTTGGGC 2097
Db 1561 CTTCACAGCAGCAGAGCACTCAAGAAAGTGTGCTGCTGAGGCTTAAAGGCTGAC 1620
QY 2098 CTTCACAGCAGCAGAGCACTCAAGAAAGTGTGCTGCTGAGGCTTAAAGGCTGAC 2157
Db 1621 TGGATATTTTGGCCCTCAGACCTTACCCCTTCCCTGACAGCAGCTGATTTTGGC 1680
QY 2158 TGGATATTTTGGCCCTCAGACCTTACCCCTTCCCTGACAGCAGCTGATTTTGGC 2217
Db 1681 ACAGAGCTCCTCACTTACTCTGCTCAGCCATCTAGAGGCACTGCACTTACTCT 1740
```


[illegible]

Db		1921	AGGTCCACGGGAAGAAGCTGGGGAAGAATGGCGCAGTCAAGTCCTTTTGCGGCAGCATGAA	1980
Oy		1981	ATCATGTAGGCTCTCC	1995
Db		1981	ATCATGTAGGCTCTCC	1995
RESULT 4				
AX278461				
LOCUS				
DEFINITION		Sequence 1 from Patent WO0177340.		
ACCESSION		AX278461		
VERSION		AX278461.1		
KEYWORDS		GI:16605915		
SOURCE				
ORGANISM		Homo sapiens (human)		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE				
AUTHORS		Duecker,K.		
TITLE		Identification of a dual specificity phosphatase: dusp-10		
JOURNAL		Patent: WO 0177340-A 1 18-OCT-2001;		
FEATURES		MERCK PATENT GmbH (DB)		
Source		Location/Qualifiers		
		1..3059		
CDS		/organism="Homo sapiens"		
		/mol_type="unassigned DNA"		
		/db_xref="taxon:9606"		
		127..2124		
		/note="unnamed protein product"		
		/codon_start=1		
		/protein_id="CAD10545.1"		
		/db_xref="GI:16605916"		
		/translation="MAHEMIGTOIVTERILVALLLESGTEKYLIDSRPFVEYNTSHILE AININCKMKMRLOQDKVLITELIOWASGHXVDIOSOVVVYPOSSOVAISLSPCC FLTVLLGLKEKSFVSHLLAGFAEFSRCPCGLBEGKSTVLPVICISOPCLPVANIQT RLIPNLVLCQDVLNKLMEQONGIYVLNASATCPDPDI PESHFLRVAVNDSPCEK ILPLDKSVGFIEKAKASNCVCVHLCLAGISRSAITAIAVIMKRMWLSDEAYRVEK KRPIISNPFNFGOLDIYEKKIKNOTGASGPSKTLILAIEKEMNEPVPVASEGOKSE TPISPCCADATSEAGORPVHPASVPSVSVOPLLEDSPLOVALSGHLSDRILED SNLKRSFSLDIKSYSIASAMASHIFESSEDALFYVKRSTLLDGNTKLCSPPVGE LSGETPETSPPKEKESI PKKLQIARPDSDOSKMLHSVRITSSETAQRSLSPHRSS VEDNYHTSFLRGISTSQHLTKSAGLKGMSHDIIAPOTSPSLTSWVFATESHF YNSAISYGGSAYSAYSQCQLPTPCGDQVYSVRRORKSPDRADSRSWHSESPEKOPK RRCQMRFESIMSERNRSRELGVGSQSFSGMTIIEVS"		
ORIGIN				
Query Match		99.8%; Score 1991.8; DB 6; Length 3059;		
Best Local Similarity		99.9%; Pred. No. 0;		
Matches 1993; Conservative		0; Mismatches 2; Indels 0; Gaps 0;		
Oy		1	ATGCCCATGAGATGATTGGTAACCTCAATTGTTACTGAGAGCTTGCTGCTCTGGTGA	60
Db		127	ATGGCCCATGAGATGATTGGAACCTCAATTGTTACTGAGAGCTTGCTGCTGGTGA	186
Oy		61	AGTGAACGGAAAAAGTGTCTTAATTGATAGCGCGCATTTGTGGAATACAATACATCC	120
Db		187	AGTGAACGGAAAAAGTGTCTTAATTGATAGCGCGCATTTGTGGAATACAATACATCC	246
Oy		121	CACATTTTGAAGCCATTAAATATCACTGCTCCAAGCTTATGAAGCGAAGTTGACAAG	180
Db		247	CACATTTTGAAGCCATTAAATATCACTGCTCCAAGCTTATGAAGCGAAGTTGACAAG	306
Oy		181	GACAAAGTGTAAATTAACAGAGCTCATTCACACATTCAGCGGAACTAAGTGTGACATTGAT	240
Db		307	GACAAAGTGTAAATTAACAGAGCTCATTCACACATTCAGCGGAACTAAGTGTGACATTGAT	366
Oy		241	TGCAGTCAGAAAGTTGTAGTTTACGATCAAGAGCTCCCAAGATGTGGCTCTCTCTTTCA	300
Db		367	TGCAGTCAGAAAGTTGTAGTTTACGATCAAGAGCTCCCAAGATGTGGCTCTCTCTTTCA	426
Oy		301	GACTGTTTTCTCACTGACTTCTTGCGTAACTGGAGAAGGCTTCAACTCTGTTCACTTG	360

Db 427 GACGTTTCTCACTGTACTTGGGTAACGAGAAAGACCTTCAACTGATCACTG 486
Qy 361 CTGGAGGAGGAGGTTTGTCTGAGTTCTCTGTTGTTTCCCTGGGCTCTGTGAAGAAATCC 420
Db 487 CTGGAGGAGGAGGTTTGTCTGAGTTCTCTGTTGTTTCCCTGGGCTCTGTGAAGAAATCC 546
Qy 421 ACTCTAGTCCCTACCTGATTTCTCAGGCTTGTACTCTGTGCAACATTGGGCAAC 480
Db 547 ACTCTAGTCCCTACCTGATTTCTCAGGCTTGTACTCTGTGCAACATTGGGCAAC 606
Qy 481 CGAATTTCTCCCACTTATCTTGGCTGCGCAGGAGATGCTCTCAACAAGAGCGGATA 540
Db 607 GGAATTTCTCCCACTTATCTTGGCTGCGCAGGAGATGCTCTCAACAAGAGCGGATA 666
Qy 541 CAGCAAGATGGGATTTGTTATGTTTAAATGCGAGCTATACCTGTTCAAGCTGACTTT 600
Db 667 CAGCAAGATGGGATTTGTTATGTTTAAATGCGAGCAATACCTGTTCAAGCTGACTTT 726
Qy 601 ATCCCGAGTCTCAATTTCCCTGCTGCTGCTGTAATGACAGCTTTTGTGAAGAAATTTTG 660
Db 727 ATCCCGAGTCTCAATTTCCCTGCTGCTGCTGTAATGACAGCTTTTGTGAAGAAATTTTG 786
Qy 661 CCGTGGTTGACAATCAGTATTTCAATTTGAGAAAGAAAAGCTCCATGATGATGTT 720
Db 787 CCGTGGTTGACAATCAGTATTTCAATTTGAGAAAGAAAAGCTCCATGATGATGTT 846
Qy 721 CTAGTCACTGTTTACTGGGATCTCCGCTCCGCTCCGACATGCTATGCTATCATCATG 847
Db 847 CTAGTCACTGTTTACTGGGATCTCCGCTCCGCTCCGACATGCTATGCTATCATCATG 906
Qy 781 AAGAGATGAGACATGTTTATGATGAGCTTACAGATTTGTAAGAAAGAAAGCTTACT 840
Db 907 AAGAGATGAGACATGTTTATGATGAGCTTACAGATTTGTAAGAAAGAAAGCTTACT 966
Qy 841 ATATCTCCAACTTCAATTTCTGGGCAACTCTGGAATATGAGAAAGAAAGCTTACT 966
Db 967 ATATCTCCAACTTCAATTTCTGGGCAACTCTGGAATATGAGAAAGAAAGCTTACT 1026
Qy 901 CAGACTGAGCATCAGGGCCAAAGAGCAATCTCAGCTGCTGCACTGAGAAAGCAAT 960
Db 1027 CAGACTGAGCATCAGGGCCAAAGAGCAATCTCAGCTGCTGCACTGAGAAAGCAAT 1086
Qy 961 GAACCTGTCCCTGCTGCTCAGAGGGTGAAGAGAAAGGAGAGCGCCCTCAGTCAACC 1020
Db 1087 GAACCTGTCCCTGCTGCTCAGAGGGTGAAGAGAAAGGAGAGCGCCCTCAGTCAACC 1146
Qy 1021 TGTGCGGACTGTGCTGCTCAGAGGAGAGCAAGAGCGCCGCTGCAAGCGCTG 1080
Db 1147 TGTGCGGACTGTGCTGCTCAGAGGAGAGCAAGAGCGCCGCTGCAAGCGCTG 1206
Qy 1081 CCCAGCGTCCGAGCGTGCAGCGCTGCTGTAAGAGAGCGCCGCTGTAAGCGCTG 1140
Db 1207 CCCAGCGTCCGAGCGTGCAGCGCTGCTGTAAGAGAGCGCCGCTGTAAGCGCTG 1266
Qy 1141 AGTGGGCTCAGCTGCTGCGCAGAGAGGCTGGAAGACAGCAATAGCTCAGGCTTCTTC 1200
Db 1267 AGTGGGCTCAGCTGCTGCGCAGAGAGGCTGGAAGACAGCAATAGCTCAGGCTTCTTC 1326
Qy 1201 TCTCTGATATCAATCAGTTTCAATTTACAGCGAGCATGCGAGATCTTTACATGCTTC 1260
Db 1327 TCTCTGATATCAATCAGTTTCAATTTACAGCGAGCATGCGAGATCTTTACATGCTTC 1386
Qy 1261 TCCCATGAGAAATGCTTTGAAATCTACAAACTTTCATCTCTGATGGGAGCAAC 1320
Db 1387 TCCCATGAGAAATGCTTTGAAATCTACAAACTTTCATCTCTGATGGGAGCAAC 1446
Qy 1321 AAGCTATGCGAGTTTCCCTGTTTCAAGAACTATCGAGAGCAATCCGAAACAGTCTCT 1380
Db 1447 AAGCTATGCGAGTTTCCCTGTTTCAAGAACTATCGAGAGCAATCCGAAACAGTCTCT 1506
Qy 1381 GATAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440

Db 1507 GATAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1566
Qy 1441 AGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
Db 1567 AGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1626
Qy 1501 TCTTCCAGCTGATCGAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
Db 1627 TCTTCCAGCTGATCGAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1686
Qy 1561 CTTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
Db 1687 CTTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1746
Qy 1621 TCGGATATCTTGGGCTCCCGAGAGCTTACCCCTTCCCTGAGCAGAGGAGGAGGAGGAG 1680
Db 1747 TCGGATATCTTGGGCTCCCGAGAGCTTACCCCTTCCCTGAGCAGAGGAGGAGGAGGAG 1806
Qy 1681 ACAGAGTCTTCACTTCTTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
Db 1807 ACAGAGTCTTCACTTCTTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1866
Qy 1741 GCTTACGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
Db 1867 GCTTACGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1926
Qy 1801 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
Db 1927 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1986
Qy 1861 AAGCACTTTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
Db 1987 AAGCACTTTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2046
Qy 1921 AGGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
Db 2047 AGGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2106
Qy 1981 ATCATTTGAGGCTCC 1995
Db 2107 ATCATTTGAGGCTCC 2121

RESULT 5
AX441210 3496 bp DNA linear PAT 28-JUN-2002
LOCUS AX441210
DEFINITION Sequence 1 from Patent WO0226997.
ACCESSION AX441210
VERSION AX441210.1 GI:21665766
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Luche, R.M. and Wei, B.
AUTHORS Dep-16 dual-specificity phosphatase
TITLE Patent: WO 0226997-A 1 04-APR-2002;
JOURNAL Ceptlyr, Inc. (US)
FEATURES
source
1..3496
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1991.8; DB 6; Length 3496;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Db 562 ATGGCCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621

QY 61 AGTGAACGGAAGAGCTGCTAATTGATAGCCGGCATTTGTGAAATACATACATCC 120
DB 622 AGTGAACGGAAGAGCTGCTAATTGATAGCCGGCATTTGTGAAATACATACATCC 681
QY 121 CACATTTTGAAGCATTATATCACTGCTCAAGCTTATGAAGCAAGTTGCAACAG 180
DB 682 CACATTTTGAAGCATTATATCACTGCTCAAGCTTATGAAGCAAGTTGCAACAG 741
QY 181 GACAAAGTTTAAATACAGAGCTATCAGCAATTCAGCAAGCAATTAAGTTGACATTTGAT 240
DB 742 GACAAAGTTTAAATACAGAGCTATCAGCAATTCAGCAAGCAATTAAGTTGACATTTGAT 801
QY 241 TGCAGTCAGAAAGTTGATTTAGATTCAGCAAGCTCCCAAGATTTGGCTCTCTCTCA 300
DB 802 TGCAGTCAGAAAGTTGATTTAGATTCAGCAAGCTCCCAAGATTTGGCTCTCTCTCA 861
QY 301 GACTGTTTCTCACTGACTTCTGGGTAATCTGGAGAAAGACTTCAACTCTGTCACTG 360
DB 862 GACTGTTTCTCACTGACTTCTGGGTAATCTGGAGAAAGACTTCAACTCTGTCACTG 921
QY 361 CTTCAGAGTGGGTTTGTGAGATCTCTGTTTCCCTGCTGCTGTAAGAAATCC 420
DB 922 CTTCAGAGTGGGTTTGTGAGATCTCTGTTTCCCTGCTGCTGTAAGAAATCC 981
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTTACCTGTGCAATTTGGCCCAAC 480
DB 982 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTTACCTGTGCAATTTGGCCCAAC 1041
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGAGAGAGATGCTCCCAAGAGAGCTGATA 540
DB 1042 CGAATTTCTCCCAATCTTTATCTTGGCTGAGAGAGATGCTCCCAAGAGAGCTGATA 1101
QY 541 CAGCAGAAATGGATTTGTTATGTTTAAATGCGCACTATACTGTCCAAAGCTGACTTT 600
DB 1102 CAGCAGAAATGGATTTGTTATGTTTAAATGCGCACTATACTGTCCAAAGCTGACTTT 1161
QY 601 ATCCCGCAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 1162 ATCCCGCAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1221
QY 661 CCGGTTGTTGACAAATGATGATTTTCAATGAGAAAGCAAAAGCTCCCAATGATGTTT 720
DB 1222 CCGGTTGTTGACAAATGATGATTTTCAATGAGAAAGCAAAAGCTCCCAATGATGTTT 1281
QY 721 CTAGTCACTGTTTAGCTGGGATCTCCGCTCCGCACTATGCTATGCTATCATCATG 780
DB 1282 CTAGTCACTGTTTAGCTGGGATCTCCGCTCCGCACTATGCTATGCTATCATCATG 1341
QY 781 AAGAGATGACATGCTTTTATGATGAAAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 840
DB 1342 AAGAGATGACATGCTTTTATGATGAAAGCTTACAGATTTGTGAAGAAAGAAAGCTACT 1401
QY 841 ATATCTCAAACTCAATTTTCTGGGCACTCCTGAGCTATGAGAAAGATTAAGAAC 900
DB 1402 ATATCTCAAACTCAATTTTCTGGGCACTCCTGAGCTATGAGAAAGATTAAGAAC 1461
QY 901 CAGAGTGAAGATCAGGAGCAAAAGAGCAAACTCAAGCTGCTGCACTGAGAAAGCAAAAT 960
DB 1462 CAGAGTGAAGATCAGGAGCAAAAGAGCAAACTCAAGCTGCTGCACTGAGAAAGCAAAAT 1521
QY 961 GAACCTGTCCCTGTCTGTCTAGAGAGGTGAGACAAAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 1522 GAACCTGTCCCTGTCTGTCTAGAGAGGTGAGACAAAGAGAGAGAGAGAGAGAGAGAGAG 1581
QY 1021 TGTGCGAGCTGTGCTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1582 TGTGCGAGCTGTGCTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1641
QY 1081 CCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1642 CCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1701

QY 1141 AGTGGCTGACCTGTGCTCCGAGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1702 AGTGGCTGACCTGTGCTCCGAGAGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1761
QY 1201 TCTGTGATATCAATCACTTTCATATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1762 TCTGTGATATCAATCACTTTCATATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1821
QY 1261 TCTGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1822 TCTGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1881
QY 1321 AAGCTATGCAAGTTTCCCTGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1882 AAGCTATGCAAGTTTCCCTGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1941
QY 1381 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1942 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2001
QY 1441 AGCAAGCATTGATTTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 2002 AGCAAGCATTGATTTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2061
QY 1501 TCTTCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB 2062 TCTTCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2121
QY 1561 CTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB 2122 CTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2181
QY 1621 TCGGATATCTTGGGCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB 2182 TCGGATATCTTGGGCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2241
QY 1681 ACAGAGTCCCACTTCTACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
DB 2242 ACAGAGTCCCACTTCTACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2301
QY 1741 GCCTACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
DB 2302 GCCTACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2361
QY 1801 CAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
DB 2362 CAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2421
QY 1861 AAGCAGTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
DB 2422 AAGCAGTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2481
QY 1921 AGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
DB 2482 AGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2541
QY 1981 ATCATTTAGGCTTCC 1995
DB 2542 ATCATTTAGGCTTCC 2556

RESULT 6
AB052156 3521 bp mRNA linear PRI 18-OCT-2001
LOCUS
DEFINITION Homo sapiens MKP-7 mRNA for MAPK phosphatase-7, complete cds.
ACCESSION AB052156
VERSION AB052156.1 GI:13548676
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1

QY 1501 TCCTCACTGATGGAAGTGGAGCGTGGAGAGCAATTACACACAGCTTCTTTTCGCG 1560
DB 2064 TCTCCATGTCATGAAAGTGGAGCGTGGAGAGCAATTACACACAGCTTCTTTTCGCG 2123
QY 1561 CTTTCCACGACGACGACGACCTCAGCAAGTCTGTGCGCTTGAAGGCTGAGC 1620
DB 2124 CTTTCCACGACGACGACGACCTCAGCAAGTCTGTGCGCTTGAAGGCTGAGC 2183
QY 1621 TCGGATATCTTGGCCCCCGACAGCTTCTTCCCTTCCCTGACGACGCTGTAATTTTGGC 1680
DB 2184 TCGGATATCTTGGCCCCCGACAGCTTCTTCCCTTCCCTGACGACGCTGTAATTTTGGC 2243
QY 1681 ACAGAGCTCTCACTTCTAATCTGCTGACGACCTCAATGAGGAGGAGGCAATTAATCTCT 1740
DB 2244 ACAGAGCTCTCACTTCTAATCTGCTGACGACCTCAATGAGGAGGAGGCAATTAATCTCT 2303
QY 1741 GCCTACAGCTGACGACGACCTGCTGCGAGACCAAGTCTTCTGTGCGCAGGCG 1800
DB 2304 GCCTACAGCTGACGACGACCTGCTGCGAGACCAAGTCTTCTGTGCGCAGGCG 2363
QY 1801 CAGAGCCAAAGTACAGAGCTGACTGCGCGGAGGAGCTGCGATGAGAGAGCCCTTTGAA 1860
DB 2364 CAGAGCCAAAGTACAGAGCTGACTGCGCGGAGGAGCTGCGATGAGAGAGCCCTTTGAA 2423
QY 1861 AACGAGTTTAAAGCGAAGCTGCAATGGAATTTGAGAGAGCAATGATGAGAAC 1920
DB 2424 AACGAGTTTAAAGCGAAGCTGCAATGGAATTTGAGAGAGCAATGATGAGAAC 2483
QY 1921 AGGTCAAGGAGAGCTGCGGAGAGTGGAGCTGAGCTTCTTCTGCGGACAGTGA 1980
DB 2484 AGGTCAAGGAGAGCTGCGGAGAGTGGAGCTGAGCTTCTTCTGCGGACAGTGA 2543
QY 1981 ATCATTGAGGTCTCC 1995
DB 2544 ATCATTGAGGTCTCC 2558

RESULT 7
AX260340 3544 bp DNA linear PAT 26-OCT-2001
LOCUS Sequence 1 from Patent WO0173059.
DEFINITION AX260340
ACCESSION AX260340.1 GI:16509303
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Meyers, R.A.
TITLE 38692 and 21117: dual specificity phosphatase molecules and uses
JOURNAL Patent: WO 0173059-A 1 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Source Location/Qualifiers
1. 3544
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
589_2586
/note="unamed protein product"
/codon_start=1
/protein_id="CAD10218.1"
/db_xref="GI:16509304"
/translaction="MAHBMGTQIVTERLVALLSGNKKVLIIDSRPVEVNTSHILE
AIVNLSKMRRLQDQKMLITRQIOLISAHKVDIDCSQVYVYVSSQDVASISDC
PLVTLGKEKSRNRYHLAAGFAEFSRCPGLCEGKSTIVPTICISQPLVANTCPT
RILPNTLYGQQRVUNKEMLQNGIGVLAASNTCPKPDIPESHFLRVAVNDFCEK
ILPMLDSVDFIEKASNGCVLHVICLAGISRSATTAIAIYIMRMDMSIDEAYRVEK
KRPTISPNFNLGQLLDYEKCIKQGTGASPKSLKILHLEKNEBPVAVSEGCQGS
TPISPCADSTSEAAAGRPVHPASVPSVSVQPSLLBSPVLQALSGHLISADRLSD

SNKLRSESLIDKISVYSASMAASLHGSSESDALEYKPSYTLIDGTNLCQPSPVQE
LSEORPERSPDEKESAPIKKLQTPARSPDSORLHSVRSSSGTORSLISPLHRSQS
VEDNHTSRLPSTLSOHLTSAGLKGMSHDLAPOTSPILSSVYFATRESHF
YASASITGGASAYSAYSQGLPTCCDQVYSVRROKPSRPARSRSMHESFPEKQFK
RRSCMERSESTMSNRSEILKVSQSSFSGSMELIIVS"

Query Match 99.8%; Score 1991.8; DB 6; Length 3544;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGAACTCAATTTTACTGAGAGTGGCTGCTGAGAA 60
DB 589 ATGGCCCATGAGATGATGAACTCAATTTTACTGAGAGTGGCTGCTGAGAA 648
QY 61 AGTGAACGAAAAGTGTCTGAATTTGATACCGGCCATTGTGGAATCAATACATCC 120
DB 649 AGTGAACGAAAAGTGTCTGAATTTGATACCGGCCATTGTGGAATCAATACATCC 708
QY 121 CATTTTGGAAAGCCATTAAATATCAATGCTCAAGCTTATGAGGAGGTTGCAACAG 180
DB 709 CATTTTGGAAAGCCATTAAATATCAATGCTCAAGCTTATGAGGAGGTTGCAACAG 768
QY 181 GACAAAGTGTAAATACAGAGCTCAATCCAGATTCAGCGAAATPAAGGTTGACATTGAT 240
DB 769 GACAAAGTGTAAATACAGAGCTCAATCCAGATTCAGCGAAATPAAGGTTGACATTGAT 828
QY 241 TGCAGTCAGAGGTTGATTTACATGATCAAGCTCCCAAGATGTTGCTCTCTTTCA 300
DB 829 TGCAGTCAGAGGTTGATTTACATGATCAAGCTCCCAAGATGTTGCTCTCTTTCA 888
QY 301 GACTGTTTCTCACTGATTTCTGGGTAACTGAGAAAGAGCTCACTGTGACCTG 360
DB 889 GACTGTTTCTCACTGATTTCTGGGTAACTGAGAAAGAGCTCACTGTGACCTG 948
QY 361 CTTCAGAGTGGTGTGCTGAGTTCTCGTGTGTTTCTCGGCTCTGTGAGAGAAATCC 420
DB 949 CTTCAGAGTGGTGTGCTGAGTTCTCGTGTGTTTCTCGGCTCTGTGAGAGAAATCC 1008
QY 421 ACTGATGCTTCACTGATTTCTGAGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTG 480
DB 1009 ACTGATGCTTCACTGATTTCTGAGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTG 1068
QY 481 CGAATTCCTCCCAATCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 1069 CGAATTCCTCCCAATCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1128
QY 541 CAGCAGAAATGGAATGTTGATGTTAAATGCGAGCTATACCTGTCGCAAGCTGACTTT 600
DB 1129 CAGCAGAAATGGAATGTTGATGTTAAATGCGAGCTATACCTGTCGCAAGCTGACTTT 1188
QY 601 ATCCCGAGCTCAATTTCTGCGGTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 660
DB 1189 ATCCCGAGCTCAATTTCTGCGGTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 1248
QY 661 CCGTGTGGAACAATACATGATTTTCAATGAGAAAGCAAGCCCAATGATGATGTT 720
DB 1249 CCGTGTGGAACAACAATGATTTTCAATGAGAAAGCAAGCCCAATGATGATGTT 1308
QY 721 CTAGTCACTGTTAGCTGGAGCTCCGCTCGGCAACATGCTATCGCTTACATCATG 780
DB 1309 CTAGTCACTGTTAGCTGGAGCTCCGCTCGGCAACATGCTATCGCTTACATCATG 1368
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGAAAGCTTACT 840
DB 1369 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGAAAGCTTACT 1428
QY 841 AATATCCCAACTTAAATTTTCTGGGCAACCTGGAATATGAGAAAGAAATTAAGAC 900
DB 1429 AATATCCCAACTTAAATTTTCTGGGCAACCTGGAATATGAGAAAGAAATTAAGAC 1488
QY 901 CAGACTGAGCATTCAGGCGCAAGAGCAACTCAAGCTGCTGACCTGAGAGAACCAAT 960


```
Db      1489 CAGACTGAGCATCAGGCGCAAGCAAACTCAAGCTGCTGCACTTGAGAAAGCCAAAT 1548
Qy      961  GAACCTGTCCTGCTGTCCTGAGAGGTGAGACAGAAAGCGAGAGCGCCCTCACTCCACC 1020
Db      1549 GAACCTGTCCTGCTGTCCTGAGAGGTGAGACAGAAAGCGAGAGCGCCCTCACTCCACC 1608
Qy      1021 TGTGCCGACTCTGCTTACCTCAGAGGCAAGCAAAAGCCGTGCATCCCGCAGCGTG 1080
Db      1609 TGTGCCGACTCTGCTTACCTCAGAGGCAAGCAAAAGCCGTGCATCCCGCAGCGTG 1668
Qy      1081 CCCAGCGTCCCGACCGTGCAGCGCTCCCTGTAGAGGACAGCCCGCTGTACAGGCGCTC 1140
Db      1669 CCCAGCGTCCCGACCGTGCAGCGCTCCCTGTAGAGGACAGCCCGCTGTACAGGCGCTC 1728
Qy      1141 AGTGGGCTGACCTGTCCGACAGAGGCTGAAAGACAGCAATTAAGTCAAGCGTTCCTTC 1200
Db      1729 AGTGGGCTGACCTGTCCGACAGAGGCTGAAAGACAGCAATTAAGTCAAGCGTTCCTTC 1788
Qy      1201 TCTCTGATATCAAAATCAGTTTCAATTCAGCCAGCATGCGCATCTTACATGAGCTTC 1260
Db      1789 TCTCTGATATCAAAATCAGTTTCAATTCAGCCAGCATGCGCATCTTACATGAGCTTC 1848
Qy      1261 TCCCTATCAGAAAGTGTGGAATCTCAAACTTCCACTTACTCTGATGAGCAAC 1320
Db      1849 TCCCTATCAGAAAGTGTGGAATCTCAAACTTCCACTTACTCTGATGAGCAAC 1908
Qy      1321 AAGCTATGCCAGTTCTCCCTGTTTCAAGAACTATCCGACAGACTCCGAAACCAAGTCT 1380
Db      1909 AAGCTATGCCAGTTCTCCCTGTTTCAAGAACTATCCGACAGACTCCGAAACCAAGTCT 1968
Qy      1381 GATTAAGAGAGAGCCAGCATCCCGAAGAGCTGACAGCCGCAAGGCTTCAGACAGCCAG 1440
Db      1969 GATTAAGAGAGAGCCAGCATCCCGAAGAGCTGACAGCCGCAAGGCTTCAGACAGCCAG 2028
Qy      1441 AGCAAGCAGTTGATTCGGTTCAGAAACAGCAGAGAGGACCGCCAGAGGTCCTTTTA 1500
Db      2029 AGCAAGCAGTTGATTCGGTTCAGAAACAGCAGAGAGGACCGCCAGAGGTCCTTTTA 2088
Qy      1501 TCTCCACTGTCATGAAATGGAGAGCTGAGAGCAATTAACAACAGGCTTCCTTTTGGGC 1560
Db      2089 TCTCCACTGTCATGAAATGGAGAGCTGAGAGCAATTAACAACAGGCTTCCTTTTGGGC 2148
Qy      1561 CTTTCCACAGCAGAGCAGCACTCAAGAGTGTGCTGGCTTGGGCTTAAAGGCTGGCAC 1620
Db      2149 CTTTCCACAGCAGAGCAGCACTCAAGAGTGTGCTGGGCTTAAAGGCTGGCAC 2208
Qy      1621 TGGATATCTTGGCCCCCAGACTCTACCCCTTCCTGACCAAGCAGTGGATTTTGGC 1680
Db      2209 TGGATATCTTGGCCCCCAGACTCTACCCCTTCCTGACCAAGCAGTGGATTTTGGC 2268
Qy      1681 AAGAGTCTCTCACTTTCTACTCTGCTCAGCAGCATCTACGAGAGCAGTGGCAATTTCT 1740
Db      2269 AAGAGTCTCTCACTTTCTACTCTGCTCAGCAGCATCTACGAGAGCAGTGGCAATTTCT 2328
Qy      1741 GCTACAGTGCAGCAGCTGCGCACTTGGCGGAGCAAGCTATTCTGTGCGGAGGCGG 1800
Db      2329 GCTACAGTGCAGCAGCTGCGCACTTGGCGGAGCAAGCTATTCTGTGCGGAGGCGG 2388
Qy      1801 CAGAAAGCAGTGCAGAGTGAAGTGTGCGGCGGAGCTGGCATGAAGAGAGCCCTTTGAA 1860
Db      2389 CAGAAAGCAGTGCAGAGTGAAGTGTGCGGCGGAGCTGGCATGAAGAGAGCCCTTTGAA 2448
Qy      1861 AAGCAGTTTAAAGCAGAGTGCCTAAATGGAATTTGAGAGAGCAGCATGTCAAGAAC 1920
Db      2449 AAGCAGTTTAAAGCAGAGTGCCTAAATGGAATTTGAGAGAGCAGCATGTCAAGAAC 2508
Qy      1921 AGGTCAACGGAAGAGCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTGGGAGAGATGGA 1980
Db      2509 AGGTCAACGGAAGAGCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTGGGAGAGATGGA 2568
Qy      1981 ATCATTTAGGCTTCC 1995
Db      2569 ATCATTTAGGCTTCC 2583
```

```
RESULT 8
AF506796
LOCUS
DEFINITION
Homo sapiens dual specificity phosphatase 16 (DUSP16) mRNA.
ACCESSION
AF506796
VERSION
AF506796.1 GI:25573087
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 3566)
Hoorneart, I., Marynen, P., Goris, J., Scioc, R. and Baens, M.
MPK phosphatase DUSP16/MKP-7, a candidate tumor suppressor for
chromosome region 12p12-13, reduces BCR-ABL-induced transformation
Oncogene 22 (49), 7728-7736 (2003)
14586399
2 (bases 1 to 3566)
Hoorneart, I., Marynen, P. and Baens, M.
Direct Submission
Submitted (26-APR-2002) Department for Human Genetics-Flanders
Interuniversity Institute for Biotechnology (VIB), Katholieke
Universiteit Leuven, Herestraat 49, Leuven B-3000, Belgium
FEATURES
Source
1. .3566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p13-p12"
/tissue_type="bone marrow"
1. .3566
/gene="DUSP16"
/feature="synonym: MKP7"
633. .2630
/gene="DUSP16"
/EC_number="3.1.3.48"
/EC_number="3.1.3.16"
/codon_start=1
/product="dual specificity phosphatase 16"
/protein_id="AAN75120.1"
/db_xref="GI:25573087"
/translation="MAHMEICTQVTERLVALLSSGTRKVLIDSRPFVEYNTSHILE
AININCSKLMKRRLQDDVLTTELQHSARKHVIDCSQKVVYVDSQDVASISDC
FUTVLGLKLEKSFNSVHLAAGFAEFRCFPGLCBGKSTLVPTCISQCLPVAHIGPT
RLPMLYLGCQRDVNLKELMOONGIGVLANSNTPKPDPIFESHFLRPVNDSECEK
ILPMLDKSVDFLEKAKNSGCVLVHCLAGISRSATLIAVIMKMDMSLDAARFVKE
KRPSTPNFNLGOLDYERKIKNOTGASGSRKTLILHLEKPEHPVPAVSSEGGKXB
TLPSPRCDAATSEMAAGORPVHPASVBSVPSVQSLDESLVVALGLHLSARLSD
SNKLRSTSLDIKSVYSASMAASLHGSSEDLKRYKSTTLDGNTKLCQFSPVB
LSBQTPERSPKREKSLIPKLTQIARPSDSOSKRLHSVTSSTGTAOSLSPLHRSQS
VEDNHTSFLFLGISTSOQLTKSAGLKGWHSIDILAPQISTPSLTSTWTFATESSHF
YSASLIVGSGASVAYSQSLPTCGDQVYVSVRQKPSDRADSRSHWERSPEFKQPK
RSCQMEFGESIMSENRREHIGKVGSQSPSGSMELIIVS"
ORIGIN
Query Match 99.8%; Score 1991.8; DB 9; Length 3566;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy      1 ATGGCCATGAGATGATGAACTCAATTTGTTCTGAGAGGTGGTGGCTGCTGCGA 60
Db      633 ATGGCCATGAGATGATGAACTCAATTTGTTCTGAGAGGTGGTGGCTGCTGCGA 692
Qy      61 AGTGAACGGAAGAGTCTGCTAATTGATAGCCGCGCATTTTGTGAATCAATACATCC 120
Db      693 AGTGAACGGAAGAGTCTGCTAATTGATAGCCGCGCATTTTGTGAATCAATACATCC 752
Qy      121 CACATTTGGAAGCATTAAATCACTGCTCCAGCTTATGAAGCAGAGGTTGCAACAG 180
```

Db 753 CACATTTGGAGCCATTAAATCACTGCTCCAGCTTATGAAGGAGTTGCAACAG 812
 QY 181 GAGAAAGTGTATTAACAGAGCTCATCCAGCATTTCCAGGAAACATAAGTTGACATTGAT 240
 Db 813 GAGAAAGTGTATTAACAGAGCTCATCCAGCATTTCCAGGAAACATAAGTTGACATTGAT 872
 QY 241 TGCAGTCAGAAAGGTTGATTTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 300
 Db 873 TGCAGTCAGAAAGGTTGATTTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 932
 QY 301 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAAAGACTTCACTCTGTTCACTG 360
 Db 933 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAAAGACTTCACTCTGTTCACTG 992
 QY 361 CTTCGAGTGGGTTTGTGAGTTCTCTGCTGTTTCCCTGCTCTGTGAGAAAGAAATCC 420
 Db 993 CTTCGAGTGGGTTTGTGAGTTCTCTGCTGTTTCCCTGCTCTGTGAGAAAGAAATCC 1052
 QY 421 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGCTTACCTGTGCAACATTTGGGCCAAC 480
 Db 1053 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGCTTACCTGTGCAACATTTGGGCCAAC 1112
 QY 481 CGAATTTCTTCCAAATCTTATCTTGTGCTGCGAGCGAGATGTCTCAACAAGAGCTGATA 540
 Db 1113 CGAATTTCTTCCAAATCTTATCTTGTGCTGCGAGCGAGATGTCTCAACAAGAGCTGATA 1172
 QY 541 CAGCAGAAATGGGATGGTTATGTGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 600
 Db 1173 CAGCAGAAATGGGATGGTTATGTGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 1232
 QY 601 ATCCCGAGCTCATATTTCTGCGGTGCTCTGTAATGACAGACTTTTGTGAGAAATTTTG 660
 Db 1233 ATCCCGAGCTCATATTTCTGCGGTGCTCTGTAATGACAGACTTTTGTGAGAAATTTTG 1292
 QY 661 CCGTGTGTGAGCAAAATCAGTAAATTTCAATGAGAAAGCAAAAGCCTTCAATGATGTGTT 720
 Db 1293 CCGTGTGTGAGCAAAATCAGTAAATTTCAATGAGAAAGCAAAAGCCTTCAATGATGTGTT 1352
 QY 721 CTAGTGCATGTTTAAAGTGGATCTCCCGCTCCGCCACATCCGTATCCGCTTCAATCATG 780
 Db 1353 CTAGTGCATGTTTAAAGTGGATCTCCCGCTCCGCCACATCCGTATCCGCTTCAATCATG 1412
 QY 781 AAGAGATGAGCAATGCTTATAGATGAAGCTTACAGATTTGTGAGAAAGAAAGACCTACT 840
 Db 1413 AAGAGATGAGCAATGCTTATAGATGAAGCTTACAGATTTGTGAGAAAGAAAGACCTACT 1472
 QY 841 ATATCTTCAAACTTCAATTTTCTGCGCAACTCTGTGACTATGAGAAAGATTAAGAAC 900
 Db 1473 ATATCTTCAAACTTCAATTTTCTGCGCAACTCTGTGACTATGAGAAAGATTAAGAAC 1532
 QY 901 CAGACTGAGGATCAGGGGCCAAAGAGCAAACTCAAGTGTCTGCACTGTGAGAAAGCCAAAT 960
 Db 1533 CAGACTGAGGATCAGGGGCCAAAGAGCAAACTCAAGTGTCTGCACTGTGAGAAAGCCAAAT 1592
 QY 961 GAACCTGTCCCTGCTGTCTCAGAGGGGTGAGCAGAAAGCGAGGAGCCCTCAGTCCACC 1020
 Db 1593 GAACCTGTCCCTGCTGTCTCAGAGGGGTGAGCAGAAAGCGAGGAGCCCTCAGTCCACC 1652
 QY 1021 TGTGCCGACTCTGCTACTCAGAGGAGCAGAGCAAAAGGCCCTGTGATCCCGCAGAGTG 1080
 Db 1653 TGTGCCGACTCTGCTACTCAGAGGAGCAGAGCAAAAGGCCCTGTGATCCCGCAGAGTG 1712
 QY 1081 CCGAGGCTGCCAGCGCTGAGCGCTGTCTGTGAGAGACAGCCCGCTGTGATCAGGGGCTC 1140
 Db 1713 CCGAGGCTGCCAGCGCTGAGCGCTGTCTGTGAGAGACAGCCCGCTGTGATCAGGGGCTC 1772
 QY 1141 AGTGGGCTGACCTGTCCGAGAGAGGCTGGAAGAGCAATAGCTCAAGCTTCTTC 1200
 Db 1773 AGTGGGCTGACCTGTCCGAGAGAGGCTGGAAGAGCAATAGCTCAAGCTTCTTC 1832
 QY 1201 TCTCTGATATCAAAATCAGTTTCAATATTCAGCCAGCATGCGACATCTTCAATGAGCTTC 1260
 Db 1833 TCTCTGATATCAAAATCAGTTTCAATATTCAGCCAGCATGCGACATCTTCAATGAGCTTC 1892

QY 1261 TCTCATCAGAAAGATGCTTGGAACTACTCAAACTTCCAGCTACTGATGAGCAAC 1320
 Db 1893 TCTCATCAGAAAGATGCTTGGAACTACTCAAACTTCCAGCTACTGATGAGCAAC 1952
 QY 1321 AAGCTATGCGAGTCTTCCCTGTTCAGGAATATCGAGAGAGACTCCCGAAACAGTCT 1380
 Db 1953 AAGCTATGCGAGTCTTCCCTGTTCAGGAATATCGAGAGAGACTCCCGAAACAGTCT 2012
 QY 1381 GATTAAGAGGAGGCCAGCATCCCAAGAAAGCTGACACCGCCAGGCTTTCAGACAGCAG 1440
 Db 2013 GATTAAGAGGAGGCCAGCATCCCAAGAAAGCTGACACCGCCAGGCTTTCAGACAGCAG 2072
 QY 1441 AGCAAGGATTTGATTTGCTGCTCAGAAACACAGAGAGTGGCAGCCGAGAGGTCCCTTTA 1500
 Db 2073 AGCAAGGATTTGATTTGCTGCTCAGAAACACAGAGAGTGGCAGCCGAGAGGTCCCTTTA 2132
 QY 1501 TCTCCACTGCTATGAAAGTGGAGCTGAGAGCAATTAACACACAGCTTCTTTCCG 1560
 Db 2133 TCTCCACTGCTATGAAAGTGGAGCTGAGAGCAATTAACACACAGCTTCTTTCCG 2192
 QY 1561 CTTTCCACAGCCAGCAGACCTTACGAAAGTGTCTGAGCTTGGGCTTAAAGGCTGAC 1620
 Db 2193 CTTTCCACAGCCAGCAGACCTTACGAAAGTGTCTGAGCTTGGGCTTAAAGGCTGAC 2252
 QY 1621 TGGGATATCTGGGCCCCCGAGACCTTACCCCTTCCCTGACAGAGCTGTATTTTGGC 1680
 Db 2253 TGGGATATCTGGGCCCCCGAGACCTTACCCCTTCCCTGACAGAGCTGTATTTTGGC 2312
 QY 1681 ACAGAGTCTCTACACTTCTACTCTGCTCAGCATCTAGAGAGGAGTGCAGTTACTCT 1740
 Db 2313 ACAGAGTCTCTACACTTCTACTCTGCTCAGCATCTAGAGAGGAGTGCAGTTACTCT 2372
 QY 1741 GCTTACAGCTGACAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTTGTGCGAGCG 1800
 Db 2373 GCTTACAGCTGACAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTTGTGCGAGCG 2432
 QY 1801 CAGAAAGCAAGTACAGAGCTGACTGCGCGGCGAGCTGAGATGAAGAGGCCCTTTGAA 1860
 Db 2433 CAGAAAGCAAGTACAGAGCTGACTGCGCGGCGAGCTGAGATGAAGAGGCCCTTTGAA 2492
 QY 1861 AAGCAGTTTAAACGAGAAAGCTGCCAAATGGAATTTGAGAGAGCATATGTACAGAAC 1920
 Db 2493 AAGCAGTTTAAACGAGAAAGCTGCCAAATGGAATTTGAGAGAGCATATGTACAGAAC 2552
 QY 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTACTTTTGGGAGCATGGA 1980
 Db 2553 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTACTTTTGGGAGCATGGA 2612
 QY 1981 ATCATTGAGGTCTCC 1995
 Db 2613 ATCATTGAGGTCTCC 2627

RESULT 9
 AX374994 3766 bp DNA linear PAT 01-MAR-2002
 LOCUS AX374994
 DEFINITION Sequence 17 from Patent WO0210363.
 ACCESSION AX374994
 VERSION AX374994.1 GI:19169826
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS
 1 Tang,Y.T., Elliott,V.S., Ramkumar,J., Yao,M.G., Burford,N.,
 Wang,Y.B., Stewart,E.A., Gandhi,A.R., Patterson,C., Lee,E.A.,
 Hefalia,A.J., Lu,D.A., Tribouley,C.M., Griffin,J.A., Baughn,M.R.,
 Yue,H., Warren,B.A., Nguyen,D.B. and Walla,N.K.
 Protein phosphatases
 Patent: WO 0210363-A 17 07-FEB-2002;
 Incyte Genomics, Inc. (US)

Location/Qualifiers

```
Location/Qualifiers
1..3766
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7480570CB1"
```

.....	2, inches	0; Gaps	0;
.....			

Db 1438 CAGACTGAGGATCATGAGGCGCCAAAGCAAACTCAAGTCTGTGACCTGGAGAAAGCAAAAT 1497

Qy 961 GAACTGTGCCCCGCTGTCTTCAAGAGGTGAGCAGAAAAAGGAGACGCCCCCTAGTCAACC 1020

Db 1498 GAACTGTGCCCCGCTGTCTTCAAGAGGTGAGCAGAAAAAGGAGACGCCCCCTAGTCAACC 1557

Qy 1021 TGTGCCAATCTGTGCTTACCTGAGAGGACAGAGCAAAAGGCCGTGTCAATCCCGCAGGTG 1080

Db 1558 TGTGTCCGACTGTGCTTACCTGAGAGGACAGAGCAAAAGGCCGTGTCAATCCCGCAGGTG 1617

Qy 1081 CCCAGCTGAGCCACGCTGAGAGCCGTGTGAGAGCAAGCCCGTGTGATCAAGCGCTC 1140

Db 1618 CCCAGCTGAGCCACGCTGAGAGCCGTGTGAGAGCAAGCCCGTGTGATCAAGCGCTC 1677

Qy 1141 AGTGGGCTGCACTGTGTCCGAGACAGAGCTGAGAGACACATTAAGCTCAAGCGTCTTC 1200

Db 1678 AGTGGGCTGCACTGTGTCCGAGAGAGCTGAGAGACACATTAAGCTCAAGCGTCTTC 1737

Qy 1201 TCTGTGATATCAAAATCAGTTTCAATTTACGACGACATGGCAGATCTTTCATGTGCTTC 1260

Db 1738 TCTGTGATATCAAAATCAGTTTCAATTTACGACGACATGGCAGATCTTTCATGTGCTTC 1797

Qy 1261 TCCCATATCAGAAAGATGCTTTGGAAATCTACAAACCTTTCATCTGTGATGGAGCAAC 1320

Db 1798 TCCCATATCAGAAAGATGCTTTGGAAATCTACAAACCTTTCATCTGTGATGGAGCAAC 1857

Qy 1321 AAGCTATGCGAGTCTTCCCTGTTCAGGAACTATCGAGCAGATCCGAAACAGTCTCT 1380

Db 1858 AAGCTATGCGAGTCTTCCCTGTTCAGGAACTATCGAGCAGATCCGAAACAGTCTCT 1917

Qy 1381 GATTAAGAGAGGAGCGAGATCCGCAAAACCTGACACCGCAGCTTTCAGACAGCCAG 1440

Db 1918 GATTAAGAGAGGAGCGAGATCCGCAAAACCTGACACCGCAGCTTTCAGACAGCCAG 1977

Qy 1441 AAGCAAGCCATTTGCAATTTGGTCAAGAACAGCAGCAGTGTGACCGCCAGAGGCTCTTTA 1500

Db 1978 AAGCAAGCCATTTGCAATTTGGTCAAGAACAGCAGCAGTGTGACCGCCAGAGGCTCTTTA 2037

Qy 1501 TCTCCACTGATCGAATGTGGAGACGTGGAGAGCAATTAACACACAGTCTCTTTTGCGC 1560

Db 2038 TCTCCACTGATCGAATGTGGAGACGTGGAGAGCAATTAACACACAGTCTCTTTTGCGC 2097

Qy 1561 CTTTTCACACAGCCAGAGGACCTTCACAGAAATCTGTGCGCTGTGGCTTTAAGGCTGGCAC 1620

Db 2098 CTTTTCACACAGCCAGAGGACCTTCACAGAAATCTGTGTGCGCTGTGGCTTTAAGGCTGGCAC 2157

Qy 1621 TCGAATATCTTTGGCCCCCAGACCTTCACCCCTTCCCGACACAGACCTGTATTTTGGC 1680

Db 2158 TCGAATATCTTTGGCCCCCAGACCTTCACCCCTTCCCGACACAGACCTGTATTTTGGC 2217

Qy 1681 ACAAGATCTCAACATTTCACTCTGACCTGACCCATCTACGAGGACAGTGCAGATTACTCT 1740

Db 2218 ACAAGATCTCAACATTTCACTCTGACCTGACCCATCTACGAGGACAGTGCAGATTACTCT 2277

Qy 1741 GCTTACAGCTGCAAGCCAGCTGCCCACTTTCGAGAGACCAATCTAATCTGTGGAGAGGAG 1800

Db 2278 GCTTACAGCTGCAAGCCAGCTGCCCACTTTCGAGAGACCAATCTAATCTGTGGAGAGGAG 2337

Qy 1801 CAGAAACCAAGTGA CAGAGCTGATCTGCGCGCGAGAGCTGGCATGAAGAGAGCCCTTTTGA 1860

Db 2338 CAGAAACCAAGTGA CAGAGCTGATCTGCGCGCGAGAGCTGGCATGAAGAGAGCCCTTTTGA 2397

Qy 1861 AAGCAATTTAAACCGAAGCTGCGCAATGAATTTTGAAGAGCATCATGTCAAGAGAC 1920

Db 2398 AAGCAATTTAAACCGAAGCTGCGCAATGAATTTTGAAGAGCATCATGTCAAGAGAC 2457

Qy 1921 AAGTCAAGGAGAGCTGTGGGAAAGTGGGAGTCAAGTCAAGCTTTTGGCGAGCATGAA 1980

Db 2458 AAGTCAAGGAGAGCTGTGGGAAAGTGGGAGTCAAGTCAAGCTTTTGGCGAGCATGAA 2517

Qy 1981 ATCATTTAGAGTCTCC 1995

QY 1141 AGTGGCTGACCTGTCCGAGACAGGCTGGAGACAGATAAGTCAAGCTTCCCTTC 1200
 Db 1324 AGTGGCTGACCTGTCCGAGACAGGCTGGAGACAGATAAGTCAAGCTTCCCTTC 1383
 QY 1201 TCTCTGGATATCAATATGATTTTCAATATTCAGCCAGATGGCAGATCTTACATGCTTC 1260
 Db 1384 TCTCTGGATATCAATATGATTTTCAATATTCAGCCAGATGGCAGATCTTACATGCTTC 1443
 QY 1261 TCCCATCAGAAAGTGTGGTGGAACTACTAAGAACTTCCCTCACTACTCTGGATGGAGACCAAC 1320
 Db 1444 TCCCATCAGAAAGTGTGGTGGAACTACTAAGAACTTCCCTCACTACTCTGGATGGAGACCAAC 1503
 QY 1321 AAGCTATGCAAGTCTCCCTGTTCAGAACTATCGAGAGAGATCTCCGAAACCAAGTCT 1380
 Db 1504 AAGCTATGCAAGTCTCCCTGTTCAGAACTATCGAGAGAGATCTCCGAAACCAAGTCT 1563
 QY 1381 GATTAAGAGAGAAAGCAGACATCCCAAGAGCTGAGACCCGACAGCTTTCAGACAGCCAG 1440
 Db 1564 GATTAAGAGAGAAAGCAGACATCCCAAGAGCTGAGACCCGACAGCTTTCAGACAGCCAG 1623
 QY 1441 AGCAAGGATTTGCTTGGCTCAGAAACAGACAGAGTGGCAACCGCCAGAGAGTCCCTTTTA 1500
 Db 1624 AGCAAGGATTTGCTTGGCTCAGAAACAGACAGAGTGGCAACCGCCAGAGAGTCCCTTTTA 1583
 QY 1501 TCTTCACTGATCGAAGTGGAGGCTGAGAGACATTAACACACAGCTTCTTTTGGG 1560
 Db 1684 TCTTCACTGATCGAAGTGGAGGCTGAGAGACATTAACACACAGCTTCTTTTGGG 1743
 QY 1561 CTTTCCACAGCCAGACAGACCTTACAGAGTCTGCTGGCTTGGGCTTTAAAGGCTGGCAC 1620
 Db 1744 CTTTCCACAGCCAGACAGACCTTACAGAGTCTGCTGGCTTGGGCTTTAAAGGCTGGCAC 1803
 QY 1621 TCGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACACAGACTGATATTTGGC 1680
 Db 1804 TCGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACACAGACTGATATTTGGC 1863
 QY 1681 ACAGAGTCTTCACTACTTCTGCTGCTGACAGCCATCTACAGAGGAGTGGCAGTTACTCT 1740
 Db 1864 ACAGAGTCTTCACTACTTCTGCTGCTGACAGCCATCTACAGAGGAGTGGCAGTTACTCT 1923
 QY 1741 GCCTACAGCTGACAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTTCTGTGCGAGCGG 1800
 Db 1924 GCCTACAGCTGACAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTTCTGTGCGAGCGG 1983
 QY 1801 CAGAAAGCAAGTGAAGAGTGAAGTCTGCGCGGAGCTGGCATGAAGAGACCCCTTTGAA 1860
 Db 1984 CAGAAAGCAAGTGAAGAGTGAAGTCTGCGCGGAGCTGGCATGAAGAGACCCCTTTGAA 2043
 QY 1861 AAGCAAGTTTAAAGCAGAGAGTGGCAATGGAATTTGAGAGAGCATCATGTCAAGAGAC 1920
 Db 2044 AAGCAAGTTTAAAGCAGAGAGTGGCAATGGAATTTGAGAGAGCATCATGTCAAGAGAC 2103
 QY 1921 AGGTCAAGGAGAGAGTGGGAGAAAGTGGGAGTCAAGTCTTCTTGGGAGAGATGAA 1980
 Db 2104 AGGTCAAGGAGAGAGTGGGAGAAAGTGGGAGTCAAGTCTTCTTGGGAGAGATGAA 2163
 QY 1981 ATCATTGAGGCTTCC 1995
 Db 2164 ATCATTGAGGCTTCC 2178

AUTHORS Nagase, T., Kikuno, R., Hattori, A., Kondo, Y., Okumura, K. and Ohara, O.
 TITLE Prediction of the coding sequences of unidentified human genes.
 JOURNAL XIX. The complete sequences of 100 new cDNA clones from brain which
 MEDLINE code for large proteins in vitro
 PUBMED DNA Res. 7 (6), 347-355 (2000)
 REFERENCE 21082932
 11214970
 2 (bases 1 to 4790)
 AUTHORS Ohara, O., Nagase, T. and Kikuno, R.
 TITLE Direct Submission
 JOURNAL Submitted (22-NOV-2000) Osamu Ohara, Kazuo DNA Research Institute,
 Department of Human Gene Research; 1532-3, Yana, Kiearazu, Chiba
 292-0812, Japan (E-mail: cdnainfo@kazuo.or.jp,
 URL: http://www.kazuo.or.jp/huge, Tel: 81-438-52-3913,
 Fax: 81-438-52-3914)

FEATURES
 source
 1. .4790
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="fj15353"
 /note="vector:pluescriptII SK plus"
 1. .4790
 /gene="KIAA1700"
 <109. .2181
 /gene="KIAA1700"
 /note="Start codon is not identified."

CDS
 /codon_start=1
 /product="KIAA1700 protein"
 /protein_id="BAB21791.1"
 /db_xref="GI:12697945"
 /translation="AFEHHTTSSMOIKRGGKRGILVMAHEMIGTOIVTERLVLL
 ESGTEKVLIDSPRYEYNTSHLEAININCSLTMRRLQODVLTTELQSHAKHY
 DIDCSQKVVVDSDVDVSLSDCLVLVLGLKRSFNSHLIAGAPAFSPFEL
 CEKSTLVPVTCASQPLPVANIGPTILPLNLVLCGRDVLNKLMOQNGIGVLANSL
 TCEKDPDIPESHPLRPVNDSPCEKILPMLDSDVDFIERAKSNGCVLVHCLAGISRS
 ATAIAYIMKRMDSIDEAYRFYKEKRPISPFNPLGQLDYEKILKNOTGASGPS
 KLKLHLKPNPVPVPSVSGGKSETRPLSPKESLSTSEAGORPVHPASVPSVQ
 PSLEBSPVLQALSGHLSADRLDENKLRKSDLIKSVYSASMAASLHSGSSSD
 ALBYAPSTLLDGTNLCQSPVQELSEQTPETSPKEASITFKLQTAHPSPQSQR
 LHSVRTSSSGTQRLSLPLHRSQVSDVYHSLFLGLSTSQHLLTKSAGLKGWHS
 DILAPQSTPSLTSWYFATESHPFASAIYEGSAYSGVSCQLPTCDQVYSVR
 RQKPSRADRSRMSHESPFKQFKRSCOMEPGESIMENRREBELKVGVSQSSFSG
 SMEIIEVS"

ORIGIN
 Query Match 99.8%; Score 1991.8; DB 9; Length 4790;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATTTGAATCTCAATTTGTACTGAGAGTGGTGGCTTCTGCTGAA 60
 Db 184 ATGGCCCATGAGATGATTTGAATCTCAATTTGTACTGAGAGTGGTGGCTTCTGCTGAA 243
 QY 61 AGTGAACGGAAGAAAGTGTCTAATTTGATGATGAGCGGCAATTTGGAATCAATACATCC 120
 Db 244 AGTGAACGGAAGAAAGTGTCTAATTTGATGATGAGCGGCAATTTGGAATCAATACATCC 303
 QY 121 CACATTTTGGAGCCATTATATCACTGCTCCAGCTTATGAGAGGAGTGGCAACAG 180
 Db 304 CACATTTTGGAGCCATTATATCACTGCTCCAGCTTATGAGAGGAGTGGCAACAG 363
 QY 181 GACAAAGTGTATTTACAGAGTCAATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 240
 Db 364 GACAAAGTGTATTTACAGAGTCAATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 423
 QY 241 TGCAGTCAGAGGTTTATGATTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 360
 Db 424 TGCAGTCAGAGGTTTATGATTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 483
 QY 301 GACTGTTTCTCACTACTTCTGGTAAACGTGAGAAAGCTTCAACTCTGTTCACCTG 360
 Db 484 GACTGTTTCTCACTACTTCTGGTAAACGTGAGAAAGCTTCAACTCTGTTCACCTG 543

QY 361 CTTGACAGTGGGTTTCTGATGTTCTGTGTTTCCCTGCGCTCTGTGAGAAATCC 420
DB 544 CTTGACAGTGGGTTTCTGATGTTCTGTGTTTCCCTGCGCTCTGTGAGAAATCC 603
QY 421 ACTCTAGTCCCTGATCTGATGTTCTGAGGCTGCTTACCTGTGCGAACATTTGGGCAAC 480
DB 604 ACTCTAGTCCCTGATCTGATGTTCTGAGGCTGCTTACCTGTGCGAACATTTGGGCAAC 663
QY 481 CGAATTTCTTCCCAATCTTATCTTGGGCGCGAGATGTCCTCAACAAGAGCTGATG 540
DB 664 CGAATTTCTTCCCAATCTTATCTTGGGCGCGAGATGTCCTCAACAAGAGCTGATG 723
QY 541 CAGCGAATGGGATTTGGTATGTTAAATGCAAGCTATACCTGTGCAAGGCTGATG 600
DB 724 CAGCGAATGGGATTTGGTATGTTAAATGCAAGCTATACCTGTGCAAGGCTGATG 783
QY 601 ATCCCGAGTCTCAATTCCTGCGTGTGCGTGAATGACAGCTTTTGTGAGAAATTTTG 660
DB 784 ATCCCGAGTCTCAATTCCTGCGTGTGCGTGAATGACAGCTTTTGTGAGAAATTTTG 843
QY 661 CCGGTGTGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
DB 844 CCGGTGTGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 903
QY 721 CTAGTCACTGTTTAACTGAGATCTCCGCTCCGCAATGCTTATGCTGATCATCATG 780
DB 904 CTAGTCACTGTTTAACTGAGATCTCCGCTCCGCAATGCTTATGCTGATCATCATG 963
QY 781 AAGAGATGAGACATGCTTTAGATGATGATGATGATGATGATGATGATGATGATG 840
DB 964 AAGAGATGAGACATGCTTTAGATGATGATGATGATGATGATGATGATGATGATG 1023
QY 841 AATCTCCAACTTCAATTTTCTGCGCAATCTCTGCACTATGAGAAAGAAAGAAAG 900
DB 1024 AATCTCCAACTTCAATTTTCTGCGCAATCTCTGCACTATGAGAAAGAAAGAAAG 1083
QY 901 CAGACTGAGATCAAGGCGCAAGCAAACTCAAGCTGTGCACTGTGAGAAAGAAAG 960
DB 1084 CAGACTGAGATCAAGGCGCAAGCAAACTCAAGCTGTGCACTGTGAGAAAGAAAG 1143
QY 961 GAACTGTCTCTGCTGTCTGAGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 1144 GAACTGTCTCTGCTGTCTGAGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1203
QY 1021 TGTGCGCACTGTGCTGATCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1204 TGTGCGCACTGTGCTGATCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1263
QY 1081 CCCAGGCTGCGCAGCGCTGAGCGCTGTGTTAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1264 CCCAGGCTGCGCAGCGCTGAGCGCTGTGTTAGAGAGAGAGAGAGAGAGAGAGAG 1323
QY 1141 AGTGGCTGCACTGTGCGCAGAGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1324 AGTGGCTGCACTGTGCGCAGAGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAG 1383
QY 1201 TCTCTGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1260
DB 1384 TCTCTGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1443
QY 1261 TCTCTGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1320
DB 1444 TCTCTGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1503
QY 1321 AAGCTATGCGAGTCTTCCCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1504 AAGCTATGCGAGTCTTCCCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1563
QY 1381 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1564 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1623

QY 1441 AGCAAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
DB 1624 AGCAAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1683
QY 1501 TCTCCACTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB 1684 TCTCCACTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1743
QY 1561 CTTTCCACGAGCAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 1620
DB 1744 CTTTCCACGAGCAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 1803
QY 1621 TGGGATATCTTGGGCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB 1804 TGGGATATCTTGGGCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1863
QY 1681 ACAGAGTCTGACATCTTCTGATCTGATCTGATCTGATCTGATCTGATCT 1740
DB 1864 ACAGAGTCTGACATCTTCTGATCTGATCTGATCTGATCTGATCTGATCT 1923
QY 1741 GCTTACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
DB 1924 GCTTACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1983
QY 1801 CAGAGGCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
DB 1984 CAGAGGCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2043
QY 1861 AAGAGTTTAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
DB 2044 AAGAGTTTAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2103
QY 1921 AGGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
DB 2104 AGGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2163
QY 1981 ATCATTTAGAGTCTCC 1995
DB 2164 ATCATTTAGAGTCTCC 2178

RESULT 13
AXI80875 2732 bp DNA linear PAT 06-AUG-2001
LOCUS AXI80875
DEFINITION Sequence 2 from Patent WO0146394.
ACCESSION AXI80875
VERSION AXI80875.1 GI:15132703
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Ploeman, G.D., Martinez, R., Whyte, D., Manning, G., Sudarshan, S.,
Hill, R.J., and Flanagan, P.,
Mammalian protein phosphatases
Patent: WO 0146394-A 2 28-JUN-2001;
Sugen, Inc. (US)

FEATURES
source
1..2732
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.8%; Score 1990.2; DB 6; Length 2732;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATGAGATGATGATGATGATGATGATGATGATGATG 60
DB 538 ATGGCCCATGAGATGATGAGATGATGATGATGATGATGATGATGATG 597

QY 61 AGTGAACGGAAAAAGTGTCTAATTGATGACCGGCACTTTGTTGGAATACAATACATCC 120
DB 598 AGTGAACGGAAAAAGTGTCTAATTGATGACCGGCACTTTGTTGGAATACAATACATCC 657
QY 121 CACATTTTGGAGCCATTATATATCACTGTCTCCAGGTTATGAAGGAGTTGCAACAG 180
DB 658 CACATTTTGGAGCCATTATATCACTGTCTCCAGGTTATGAAGGAGTTGCAACAG 717
QY 181 GACAAAGTGTATATACAGAGCTATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 240
DB 718 GACAAAGTGTATATACAGAGCTATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 777
QY 241 TGCAGTCAGAAAGTTGATTTAGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 300
DB 778 TGCAGTCAGAAAGTTGATTTAGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 837
QY 301 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAGAGCTTCACTCTGTTCACTG 360
DB 838 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAGAGCTTCACTCTGTTCACTG 897
QY 361 CTTCAGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGCTCTGTAAGAGAAATCC 420
DB 898 CTTCAGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGCTCTGTAAGAGAAATCC 957
QY 421 ACTCTAGTCCCTCACTGCACTTCTCAAGCTTGTGCTTACCTGTTGCAACATTTGGGCAACC 480
DB 958 ACTCTAGTCCCTCACTGCACTTCTCAAGCTTGTGCTTACCTGTTGCAACATTTGGGCAACC 1017
QY 481 CGAATTTCTTCCCAATCTTTATCTTGTGCTGACAGAGAGATGCTCTCAACAGAGCTGATA 540
DB 1018 CGAATTTCTTCCCAATCTTTATCTTGTGCTGACAGAGAGATGCTCTCAACAGAGCTGATA 1077
QY 541 CAGAGAAATGGAGTTGGTTATGTGTTAAATGCGAGCTATCTCTGTCCAAAGCTGACTT 600
DB 1078 CAGAGAAATGGAGTTGGTTATGTGTTAAATGCGAGCTATCTCTGTCCAAAGCTGACTT 1137
QY 601 ATCCCGAGTCTAATTTCCGCGTGTGCTGTAAGTGAAGAGCTTTGTGGAATAATTTTG 660
DB 1138 ATCCCGAGTCTAATTTCCGCGTGTGCTGTAAGTGAAGAGCTTTGTGGAATAATTTTG 1197
QY 661 CCGTGTGTGACAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCTTCAATGATGTGT 720
DB 1198 CCGTGTGTGACAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCTTCAATGATGTGT 1257
QY 721 CTATGCACTGTTTATGCTGGATCTCCCGCTCCGCCACATGCTATGCTATCATCAG 780
DB 1258 CTATGCACTGTTTATGCTGGATCTCCCGCTCCGCCACATGCTATGCTATCATCAG 1317
QY 781 AAGAGATGAGCAATGCTTTAGATGAAGCTTACAGATTGTGAAGAAAGAAAGCACTACT 840
DB 1318 AAGAGATGAGCAATGCTTTAGATGAAGCTTACAGATTGTGAAGAAAGAAAGCACTACT 1377
QY 841 ATATCTCAAACTTCAATTTTCTGGGCCAATCTCTGCACTATGAGAAAGATTAAGAAC 900
DB 1378 ATATCTCAAACTTCAATTTTCTGGGCCAATCTCTGCACTATGAGAAAGATTAAGAAC 1437
QY 901 CAGACTGGAGCATCAGGGGCCAAAGAGCAATCTCAAGCTGTGCACTGTGAAGACCAAT 960
DB 1438 CAGACTGGAGCATCAGGGGCCAAAGAGCAATCTCAAGCTGTGCACTGTGAAGACCAAT 1497
QY 961 GAACCTGTCTCTGTCTCAGAGGGTGAAGAAAGCGAAGCGCCCTCACTCCACCC 1020
DB 1498 GAACCTGTCTCTGTCTCAGAGGGTGAAGAAAGCGAAGCGCCCTCACTCCACCC 1557
QY 1021 TGTGCGCACTCTGCTACTCAGAGGAGCAGAGCAAAAGGCCGTGTGATCCCGCAGAGGTG 1080
DB 1558 TGTGCGCACTCTGCTACTCAGAGGAGCAGAGCAAAAGGCCGTGTGATCCCGCAGAGGTG 1617
QY 1081 CCGAGGCTGCCAGGCTGACGCGTCTGTTAGAGAGACGCCGCTGTGACAGGCGCTC 1140
DB 1618 CCGAGGCTGCCAGGCTGACGCGTCTGTTAGAGAGACGCCGCTGTGACAGGCGCTC 1677
QY 1141 AGTGGGCTGACCTGTCTCGAGAGAGGCTGGAAGAGCAATTAAGCTCAAGCGTTCTTC 1200

DB 1678 AGTGGGCTGACCTGTCTCGAGAGAGGCTGGAAGAGCAAGAAATTAAGCTCAAGCTTCCTTC 1737
QY 1201 TCTTGTGATATCAATCACTTTCATATTTAGGCCAGATGAGCAGATCTTACATGCTTC 1260
DB 1738 TCTTGTGATATCAATCACTTTCATATTTAGGCCAGATGAGCAGATCTTACATGCTTC 1797
QY 1261 TCTTGTGATATCAATCACTTTCATATTTAGGCCAGATGAGCAGATCTTACATGCTTC 1320
DB 1798 TCTTGTGATATCAATCACTTTCATATTTAGGCCAGATGAGCAGATCTTACATGCTTC 1857
QY 1321 AAGCTATGCGAGTCTCTCCCTGTTCAGAACTATCGAGAGAGATCTCCGAAACAGTCTCT 1380
DB 1858 AAGCTATGCGAGTCTCTCCCTGTTCAGAACTATCGAGAGAGATCTCCGAAACAGTCTCT 1917
QY 1381 GATAGAGAGAAAGCCAGCATCTCCCAAGAAAGCTGAGACCGCCAGGCTTCAAGACCCAG 1440
DB 1918 GATAGAGAGAAAGCCAGCATCTCCCAAGAAAGCTGAGACCGCCAGGCTTCAAGACCCAG 1977
QY 1441 AGCAAGCATGCAATTCGGTCAAGAACAGCAGAGTGGGACCGCCAGAGGTCCTTTTA 1500
DB 1978 AGCAAGCATGCAATTCGGTCAAGAACAGCAGAGTGGGACCGCCAGAGGTCCTTTTA 2037
QY 1501 TCTCACTGCACTCAAGTGGAGAGCTGTGAGAGCAATTAACACACAGCTTCTTTTCGCG 1560
DB 2038 TCTCACTGCACTCAAGTGGAGAGCTGTGAGAGCAATTAACACACAGCTTCTTTTCGCG 2097
QY 1561 CTTCACACAGCAGAGACGCTCAGAAAGTGTGCTGAGGCTTGAAGGCTGAGCAC 1620
DB 2098 CTTCACACAGCAGAGACGCTCAGAAAGTGTGCTGAGGCTTGAAGGCTGAGCAC 2157
QY 1621 TGGATATCTGGGCCCCGAGACCTTACCCCTGCTGACAGCAGCTGTATTTTGGC 1680
DB 2158 TGGATATCTGGGCCCCGAGACCTTACCCCTGCTGACAGCAGCTGTATTTTGGC 2217
QY 1681 ACAGAGTCTTCACTTCTACTCTGCTCAGAGCATCTAAGAGGAGGCTGCACTTACTCT 1740
DB 2218 ACAGAGTCTTCACTTCTACTCTGCTCAGAGCATCTAAGAGGAGGCTGCACTTACTCT 2277
QY 1741 GCCTACAGCTGACAGCAGCTGCCCCACTTGGGAGACCAAGTCTATTTCTGTGCGAGCGG 1800
DB 2278 GCCTACAGCTGACAGCAGCTGCCCCACTTGGGAGACCAAGTCTATTTCTGTGCGAGCGG 2337
QY 1801 CAGAAAGCAGAGCAGAGCTGCTGCGGAGAGCTGAGCATGAAGAGCCCTTTGAA 1860
DB 2338 CAGAAAGCAGAGCAGAGCTGCTGCGGAGAGCTGAGCATGAAGAGCCCTTTGAA 2397
QY 1861 AAGCAGTTTAAAGCAGAGCTGCGCAATGGAATTTGAGAGAGCATGTGTCAGAGAAC 1920
DB 2398 AAGCAGTTTAAAGCAGAGCTGCGCAATGGAATTTGAGAGAGCATGTGTCAGAGAAC 2457
QY 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTGATGCTTTTGGGAGCATGAA 1980
DB 2458 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTGATGCTTTTGGGAGCATGAA 2517
QY 1981 ATCATTTAGAGTCTCC 1995
DB 2518 ATCATTTAGAGTCTCC 2532

RESULT 14
AX405700 3104 bp DNA linear PAT 14-JUN-2002
LOCUS DEFINITION Sequence 115 from Patent WO0222660.
AX405700
ACCESSION AX405700
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,

TITLE
Xue,A.-J., Yang,Y., Wehrman,T. and Drmanac,R.T.
Novel nucleic acids and polypeptides
Patent: WO 0222660-A 115 21-MAR-2002;
JOURNAL
HYSEQ, INC. (US)

FEATURES

Source

Location/Qualifiers

1..3104

/organism="Homo sapiens"

/mol_type="unassigned DNA"

CDS

/db_xref="taxon:9606"

159..2156

/note="unnamed protein product"

/protein_start=1

/protein_id="CAD34807.1"

/db_xref="GI:21438840"

/translat="RENTREMBL:CAD34807"

AININCSKLMKRLQODKVLITELIOHAKKHVDIDSOQKVVVDOSQVASSISC
FLTVLLGLKLEKSPNSVYHLAGFAFSCFGLCEKSTLPTICISOPCLPVANIPT
RLPMLYIGCORPDVYNKELMOQNGIGVYLNASNTCPKEDTIPESIFLRVPVDSFCEK
ILPMLDKSDYDFLEKASNGCVLVHCLNGISRSATIALAYIMKMDKSLDEAVRPIKE
KRPITLSPNFIQLIDYEKIKNOTGASGPKSLKLHLHMKPMPVPAVSREGQKSE
TPUSPCADASATSEAGORPVHPASVPSVPSVPSLEDSPLVQALSGHLSDRLSD
SNKLKRSFLDIKSVYSASMAASLHGPSSEEDALRYXKSTTLIDGNKLQSPVQJE
LSEQTPETSPDEKASIPKIKOTARPSPSOSKRLHVRISSTAGRSLSPLHRSQS
VEDNYHTSFLRGLSTSOOHLTKSAGLKGKMSHDILAPQSTPLSTLSWTFATESHF
YSASAIYGGASYSNYSQSLPTGCDYVSVRROKPSDRADRSRMSHESPEKQFK
RSCCMERGESIMENSRERELGRVGSQSSPSGMEIIEVS"

ORIGIN

Query Match 99.8%; Score 1990.2; DB 6; Length 3104;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAATCTCAATTTGTTACTGAGAGTGGCTCTGCTGGA 60
Db 159 ATGGCCCATGAGATGATGGAATCTCAATTTGTTACTGAGAGTGGCTCTGCTGGA 60
QY 61 AGTGAACGGAAGAGTCTGCTAATTGATAGCGCGCAATTGTGGAATACATTC 120
Db 219 AGTGAACGGAAGAGTCTGCTAATTGATAGCGCGCAATTGTGGAATACATTC 120
QY 121 CACATTTTGGAGCATTAATATCACTGCTCCAGCTTATGAGCGAGTTGGCAAG 180
Db 279 CACATTTTGGAGCATTAATATCACTGCTCCAGCTTATGAGCGAGTTGGCAAG 180
QY 181 GACAAAGTTTAATACAGAGTCATCCAGCAATTCAGCGAACAATAGTTGAAT 240
Db 339 GACAAAGTTTAATACAGAGTCATCCAGCAATTCAGCGAACAATAGTTGAAT 240
QY 241 TGCAGTCAGAGGTTGATTTACGATCAAGAGTCCCAAGATGTTGCTCTCTTCA 300
Db 399 TGCAGTCAGAGGTTGATTTACGATCAAGAGTCCCAAGATGTTGCTCTCTTCA 300
QY 301 GACTGTTTCTACATGTAATCTTGGGTAACTGAGGAAGAGCTTCACTCTTCA 360
Db 459 GACTGTTTCTACATGTAATCTTGGGTAACTGAGGAAGAGCTTCACTCTTCA 360
QY 459 GACTGTTTCTACATGTAATCTTGGGTAACTGAGGAAGAGCTTCACTCTTCA 360
Db 361 CTTCGAGGAGGTTGCTGAGTCTCTGTTGTTCCCTGGCCCTGTCGAGGAATTC 420
QY 519 CTTCGAGGAGGTTGCTGAGTCTCTGTTGTTCCCTGGCCCTGTCGAGGAATTC 420
Db 421 ACTCTAGTCCCTACCTGATTTCTCAGCTTGGCTTACCTGTCGCAACATTC 480
QY 579 ACTCTAGTCCCTACCTGATTTCTCAGCTTGGCTTACCTGTCGCAACATTC 480
Db 481 CGAATTTCTCCCATCTTTATCTTGGCTGCGCAGCGAGATGTCCTCAACAAGAGCTGAT 540
QY 639 CGAATTTCTCCCATCTTTATCTTGGCTGCGCAGCGAGATGTCCTCAACAAGAGCTGAT 540
Db 541 CAGCAGATGGGATGGTATGTTAAATGCGAGCTAATCTGTCGCAAGCTTCACTTT 600
QY 699 CAGCAGATGGGATGGTATGTTAAATGCGAGCTAATCTGTCGCAAGCTTCACTTT 600
Db 758 CAGCAGATGGGATGGTATGTTAAATGCGAGCTAATCTGTCGCAAGCTTCACTTT 600

QY 601 ATCCCCAGTCTCATTTCTGCGTGGCTGTGATGACAGCTTTTGAGAAATTTTG 660
Db 759 ATCCCCAGTCTCATTTCTGCGTGGCTGTGATGACAGCTTTTGAGAAATTTTG 660
QY 661 CCGTGGTGGCAATATCATGATGATTTCTTATGAGAAAGCAAAAGCTTCAATGATGTT 720
Db 819 CCGTGGTGGCAATATCATGATGATTTCTTATGAGAAAGCAAAAGCTTCAATGATGTT 720
QY 721 CTATGCACTGTTTATGCTGGATCTCCGCTCCGCACTTATGAGAAAGCAAAAGCTTCA 780
Db 879 CTATGCACTGTTTATGCTGGATCTCCGCTCCGCACTTATGAGAAAGCAAAAGCTTCA 780
QY 781 AAGAGATGGAATGCTCTTATGATGAGACTTACAGATTGTGAAAGAAAGCACTTACT 840
Db 939 AAGAGATGGAATGCTCTTATGATGAGACTTACAGATTGTGAAAGAAAGCACTTACT 840
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGCAATGAGAAAGCACTTACT 900
Db 999 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGCAATGAGAAAGCACTTACT 900
QY 901 CAGACTGAGATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGAGAGCAAT 960
Db 1059 CAGACTGAGATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGAGAGCAAT 960
QY 961 GAACCTGCTCTGCTCTGCTCAGAGGCTGAGCAAGAAAGCAAGCGCTCACTCACCC 1020
Db 1119 GAACCTGCTCTGCTCTGCTCAGAGGCTGAGCAAGAAAGCAAGCGCTCACTCACCC 1020
QY 1021 TGGCCGATCTGCTCTCTCAAGGCAAGAGCAAGAAAGCGCTGCACTGAGAGCAAT 1080
Db 1179 TGGCCGATCTGCTCTCTCAAGGCAAGAGCAAGAAAGCGCTGCACTGAGAGCAAT 1080
QY 1081 CCAAGCTGCCAGGCTGAGCGCTGCTGCTGTTAGAGAGCAAGCGCTGTTACAGCGCTC 1140
Db 1239 CCAAGCTGCCAGGCTGAGCGCTGCTGCTGTTAGAGAGCAAGCGCTGTTACAGCGCTC 1140
QY 1141 AGTGGCTGCACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298
Db 1299 AGTGGCTGCACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298
QY 1201 TCTTGAATATCAATCAATTCATTTCAATTCAGCGAGCAAGCAAGCAATTCATTCAT 1260
Db 1359 TCTTGAATATCAATCAATTCATTTCAATTCAGCGAGCAAGCAAGCAATTCATTCAT 1260
QY 1261 TCTTGAATATCAATTCATTTCAATTCAGCGAGCAAGCAAGCAATTCATTCAT 1418
Db 1419 TCTTGAATATCAATTCATTTCAATTCAGCGAGCAAGCAAGCAATTCATTCAT 1418
QY 1321 AAGCTATGCAATTCCTCCCTGTTCAAGAACTATGGAAGCACTTCCGAAACAGTCTT 1380
Db 1479 AAGCTATGCAATTCCTCCCTGTTCAAGAACTATGGAAGCACTTCCGAAACAGTCTT 1380
QY 1381 GATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1538
Db 1539 GATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1538
QY 1441 AGCAAGGATTTGATTCGCTGAGAACTCAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1599 AGCAAGGATTTGATTCGCTGAGAACTCAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 TCTCCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1568
Db 1659 TCTCCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1568
QY 1561 CTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1719 CTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 TCGGATATTTGGCCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1778
Db 1779 TCGGATATTTGGCCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1778
QY 1681 ACAGAGCTCAGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

Db 1839 ACAGAGCCCTCACACTTCTACTCTGCTCAGCCATCTACGAGGCAATGSCAATTACTCT 1898
Qy 1741 GCGTACAGCTGACGCCAGCTGCCCACTTGGCGAGACCAAGTCTATTCTGTGGCAGGCGG 1800
Db 1899 GCGTACAGCTGACGCCAGCTGCCCACTTGGCGAGACCAAGTCTATTCTGTGGCAGGCGG 1958
Qy 1801 CAGAGCGAAGTACAGAGCTGACTGCGGCGGAGCTGGATGAAGAGACCCCTTTGAA 1860
Db 1959 CAGAGCGAAGTACAGAGCTGACTGCGGCGGAGCTGGATGAAGAGACCCCTTTGAA 2018
Qy 1861 AACGAGTTTAAAGCAGAAAGCTGCCAAATGGAATTTGGAGAGACATCATGTCAAGAAAC 1920
Db 2019 AACGAGTTTAAAGCAGAAAGCTGCCAAATGGAATTTGGAGAGACATCATGTCAAGAAAC 2078
Qy 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCAAGTCTTTCGCGGAGCATGAAA 1980
Db 2079 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCAAGTCTTTCGCGGAGCATGAAA 2138
Qy 1981 ATCATGAGGTCTCC 1995
Db 2139 ATCATGAGGTCTCC 2153

RESULT 15

AX713989

LOCUS AX713989 2102 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 673 from Patent EPI293569.
ACCESSION AX713989
VERSION AX713989.1 GI:29888917

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 Isogai,T., Sugiyama,T., Otsuki,T., Makamatsu,A., Sato,H., Ishii,S.,
Yamanoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamehika,I., Seki,N., Yoshikawa,T., Otsuka,W., Nagaharti,K. and
Masuno,Y.
Full-length cDNAs
Patent: EP 1293569-A 673 19-MAR-2003;
Helix Research Institute (Jp) ; Research Association for
Biotechnology (Jp)

TITLE

JOURNAL

FEATURES
Source 1..2102
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Location/Qualifiers
1..2102

ORIGIN

Query Match 99.7%; Score 1988.6; DB 6; Length 2102;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATGGAACTCAATTTGTTACTGAGAGGTTGGTCTGCTGGAA 60
Db 56 ATGGCCCATGAGATGATGGAACTCAATTTGTTACTGAGAGGTTGGTCTGCTGGAA 115
Qy 61 AGTGAAGCGAAAGAGTGTCTAATTGATGAGCGGCGCATTTGAGAAATCAATACATCC 120
Db 116 AGTGAAGCGAAAGAGTGTCTAATTGATGAGCGGCGCATTTGAGAAATCAATACATCC 175
Qy 121 CACATTTGGAGCCATTATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 180
Db 176 CACATTTGGAGCCATTATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 235
Qy 181 GAGAAAGTGTATTAATGAGCTCATCCAGCATTTGAGCGAAACATAAGTTGACATTGAT 240
Db 236 GAGAAAGTGTATTAATGAGCTCATCCAGCATTTGAGCGAAACATAAGTTGACATTGAT 295
Qy 241 TGCACTCAGAAAGTGTATGATTCAGATCAAAAGCTCCCAAGATGTGCTCTCTCTCA 300

Db 296 TGCACTCAGAAAGTGTATGATTCAGATCAAAAGCTCCCAAGATGTGCTCTCTCTCA 355
Qy 301 GACTGTTTCTCAGTGAATCTTCTGGGTAAATGAGAAAGCTTCAATCTGTTCACTG 360
Db 356 GACTGTTTCTCAGTGAATCTTCTGGGTAAATGAGAAAGCTTCAATCTGTTCACTG 415
Qy 361 CTTCGAGAGTGGGTTTCTGAGTCTCTCGTGTGTTCCCTGGGCTCTGTGAAGAAATCC 420
Db 416 CTTCGAGAGTGGGTTTCTGAGTCTCTCGTGTGTTCCCTGGGCTCTGTGAAGAAATCC 475
Qy 421 ACTGTAGTCCCTTACCTGATATTTCTCAGCCCTGCTTACTGTTCGCAACATTTGGCCAAACC 480
Db 476 ACTGTAGTCCCTTACCTGATATTTCTCAGCCCTGCTTACTGTTCGCAACATTTGGCCAAACC 535
Qy 481 CGAATCTTCCCAATCTTTATCTTGGCTGCGAGGAGATGTCTCTCAAGAGAGCTGATA 540
Db 536 CTAAATCTTCCCAATCTTTATCTTGGCTGCGAGGAGATGTCTCTCAAGAGAGCTGATA 595
Qy 541 CAGCAGAAATGGGATGGTATGTTAAATGCCAGCTATACCTGTCCAAAGCCTGA 600
Db 596 CAGCAGAAATGGGATGGTATGTTAAATGCCAGCTATACCTGTCCAAAGCCTGA 655
Qy 601 ATCCCGAGTCTCATTTCTGCGCTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Db 656 ATCCCGAGTCTCATTTCTGCGCTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 715
Qy 661 CCGTGGTGAACAAATCAGTGAATTTCTATTGAGAAAGCAAAAGCCTCAATGATGTGT 720
Db 716 CCGTGGTGGACAAATCAGTGAATTTCTATTGAGAAAGCAAAAGCCTCAATGATGTGT 775
Qy 721 CTAGTGCATCTTTAGCTGGGATCTCCGCTCCGACATGCTATGCTTATCATCATG 780
Db 776 CTAGTGCATCTTTAGCTGGGATCTCCGCTCCGACATGCTTATGCTTATCATCATG 835
Qy 781 AAGAGAGTGAACATGCTTTAGTGAAGCTTACAGATTTGTGAAGAAAGAAAGCCTACT 840
Db 836 AAGAGAGTGAACATGCTTTAGTGAAGCTTACAGATTTGTGAAGAAAGAAAGCCTACT 895
Qy 841 AATCTCCAAACTTCAATTTTCTGGGCGAACTCTGAGCTATGAGAAAGATTAAGAAC 900
Db 896 AATCTCCAAACTTCAATTTTCTGGGCGAACTCTGAGCTATGAGAAAGATTAAGAAC 955
Qy 901 CAGACTGGAGCATCAGGGGCGAAAGCAAACTCAAGCTCTGACCTGAGAGAAAGCCAAAT 960
Db 956 CAGACTGGAGCATCAGGGGCGAAAGCAAACTCAAGCTCTGACCTGAGAGAAAGCCAAAT 1015
Qy 961 GAACCTGTCCCTGTCTCTCAGAGGGTGGACAGAAAGCGAGAGCGCCCTCAGTCCACCC 1020
Db 1016 GAACCTGTCCCTGTCTCTCAGAGGGTGGACAGAAAGCGAGAGCGCCCTCAGTCCACCC 1075
Qy 1021 TGTGCCAGCTGTCTACTCAGAGGCGAGAGCAAAAGGCCCTGTGATCCCGCAGCGTG 1080
Db 1076 TGTGCCAGCTGTCTACTCAGAGGCGAGAGCAAAAGGCCCTGTGATCCCGCAGCGTG 1135
Qy 1081 CCCAGGTGCCAGAGCTGAGCGCTGTGTTAGAGAGACAGCCCGCTGTGATAGAGGGCTC 1140
Db 1136 CCCAGGTGCCAGAGCTGAGCGCTGTGTTAGAGAGAGCGCCCGTGTATCAGGGCTC 1195
Qy 1141 AGTGGGCTGACCTGTCCGCGACAGCAGGCTGGAAGACAGCAATTAAGCTCAAGCGTTC 1200
Db 1196 AGTGGGCTGACCTGTCCGCGACAGCAGGCTGGAAGACAGCAATTAAGCTCAAGCGTTC 1255
Qy 1201 TCTTGATATCAATCAAGTTTCAATTCAGCCAGAGATGGCAGACCTTTACATGGCTTC 1260
Db 1256 TCTTGATATCAATCAAGTTTCAATTCAGCCAGAGATGGCAGACCTTTACATGGCTTC 1315
Qy 1261 TCTTGATATCAATCAAGTTTCAATTCAGCCAGAGATGGCAGACCTTTACATGGCTTC 1320
Db 1316 TCTTGATATCAATCAAGTTTCAATTCAGCCAGAGATGGCAGACCTTTACATGGCTTC 1375
Qy 1321 AAGCTATGCAAGTTCCTCCCTGTCAGAACTATGAGACAGACTCCGAAACAGTCTCT 1380
Db 1376 AAGCTATGCAAGTTCCTCCCTGTCAGAACTATGAGACAGACTCCGAAACAGTCTCT 1435

```
OY 1381 GATAGAGAGAGCCAGCATCCCAAGAAGCTGCAGACCCGCAAGCCTTGAAGACGACG 1440
    |||||
Db 1436 GATAGAGAGAGAGCCAGCATCCCAAGAAGCTGCAGACCTTGAAGACGACG 1495
    |||||
OY 1441 AGCAAGCATTTGCAATTGGTCAAGAACAGCAGCAGTGGCAACGCCAGAGGTCCCTTTTA 1500
    |||||
Db 1496 AGCAAGCATTTGCAATTGGTCAAGAACAGCAGCAGTGGCAACGCCAGAGGTCCCTTTTA 1555
    |||||
OY 1501 TCTCACTGCATCGAAGTGGAGAGGTGGAGACAATTACACACAGCTTCCCTTTTCGGC 1560
    |||||
Db 1556 TCTCACTGCATCGAAGTGGAGAGGTGGAGACAATTACACACAGCTTCCCTTTTCGGC 1615
    |||||
OY 1561 CTTTCCACCAAGCCAGCAGCACTTCACGAGTCTGCTGGCCTTGAAGGCTGGCAC 1620
    |||||
Db 1616 CTTTCCACCAAGCCAGCAGCACTTCACGAGTCTGCTGGCCTTGAAGGCTGGCAC 1675
    |||||
OY 1621 TCGGATATCTTGGGCCCCCGAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGC 1680
    |||||
Db 1676 TCGGATATCTTGGGCCCCCGAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGC 1735
    |||||
OY 1681 ACAGAGTCTTCACACTTCTACTCTGCTGCTAGCCATCTACGAGAGGCAAGTCACTTCT 1740
    |||||
Db 1736 ACAGAGTCTTCACACTTCTACTCTGCTGCTAGCCATCTACGAGAGGCAAGTCACTTCT 1795
    |||||
OY 1741 GCCTACAGCTGCAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGGCAGGCGG 1800
    |||||
Db 1796 GCCTACAGCTGCAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGGCAGGCGG 1855
    |||||
OY 1801 CAGAGCCAAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATGAGAGAGCCCTTTGAA 1860
    |||||
Db 1856 CAGAGCCAAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATGAGAGAGCCCTTTGAA 1915
    |||||
OY 1861 AAGCAATTAAAGCGCAAGACTGCGCAATGGAATTGAGAGAGCATCATGTCAAGAAAC 1920
    |||||
Db 1916 AAGCAATTAAAGCGCAAGACTGCGCAATGGAATTGAGAGAGCATCATGTCAAGAAAC 1975
    |||||
OY 1921 AGGTACGGGAAAGAGCTGGGAAAGTGGGCACTAGCTTTTGGGCAAGCAATGAAA 1980
    |||||
Db 1976 AGGTACGGGAAAGAGCTGGGAAAGTGGGCACTAGCTTTTGGGCAAGCAATGAAA 2035
    |||||
OY 1981 ATCATTTGAGGTCTCC 1995
    |||||
Db 2036 ATCATTTGAGGTCTCC 2050
    |||||
```

Search completed: June 21, 2004, 23:40:19
Job time : 12565.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 17:00:31 ; Search time 7898.16 Seconds

(without alignments)

7542.898 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995

Sequence: 1 atgagccatgagatgattg.....tggaaatcattgaggtctcc 1995

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthm:*
3: em_estln:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rnd:*
26: em_gss_png:*
27: em_gss_vil:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	896.8	45.0	1060	12	BM546940 AGENCOURT
2	834.6	41.8	883	13	BQ945892 AGENCOURT
3	827.8	41.5	881	13	BQ930140 AGENCOURT
4	809.6	40.6	859	13	BQ933499 AGENCOURT

Result No.	Score	Query Match	Length	ID	Description
5	799.8	40.1	898	13	BQ721265 AGENCOURT
6	786.2	39.4	974	13	BQ951695 AGENCOURT
7	780.2	39.1	903	12	BG482429 AGENCOURT
8	695.6	34.9	836	12	B1821804 AGENCOURT
9	660.6	33.1	862	13	BQ770036 AGENCOURT
10	641.2	32.1	920	10	BE897795 AGENCOURT
11	633	31.7	920	10	B1816954 AGENCOURT
12	632	31.7	663	13	BK479029 AGENCOURT
13	627	31.4	891	13	BQ670889 AGENCOURT
14	623.2	31.2	769	13	BU704078 AGENCOURT
15	615.4	30.8	660	12	B1917706 AGENCOURT
16	606.6	30.4	983	13	BG165188 AGENCOURT
17	603.2	30.2	752	13	BQ432082 AGENCOURT
18	592.6	29.7	715	14	CF727177 AGENCOURT
19	559.6	28.1	639	14	CF742387 AGENCOURT
20	558.2	28.0	656	14	CF532917 AGENCOURT
21	544.8	27.3	650	10	BF472046 AGENCOURT
22	541.4	27.1	602	10	AM847425 AGENCOURT
23	541.2	27.1	621	9	AV703072 AGENCOURT
24	536.8	26.9	693	9	AV701628 AGENCOURT
25	535.6	26.8	595	10	AM847426 AGENCOURT
26	527.2	26.4	790	12	B1661614 AGENCOURT
27	521	26.1	946	12	B1657528 AGENCOURT
28	490.8	24.6	743	14	CF727802 AGENCOURT
29	482.8	24.2	3325	11	AK035652 AGENCOURT
30	479.4	24.0	739	28	AZ850283 AGENCOURT
31	462.4	23.2	631	28	BH039241 AGENCOURT
32	455.4	22.8	1067	10	BF135687 AGENCOURT
33	433	21.7	792	13	BU946569 AGENCOURT
34	428.8	21.5	512	10	BF815601 AGENCOURT
35	425.8	21.3	586	14	CF104291 AGENCOURT
36	422.8	21.2	516	29	AY413655 AGENCOURT
37	414	20.8	723	13	BU708175 AGENCOURT
38	404.6	20.3	619	29	CG583869 AGENCOURT
39	399.6	20.0	463	10	BG058779 AGENCOURT
40	389.4	19.5	391	13	BK479224 AGENCOURT
41	386.4	19.4	463	10	BF809325 AGENCOURT
42	386.2	19.4	2207	11	BC038231 AGENCOURT
43	379	19.0	444	12	BM987142 AGENCOURT
44	373.2	18.7	516	29	AY413657 AGENCOURT
45	370.6	18.6	594	14	CF727694 AGENCOURT

ALIGNMENTS

RESULT 1
BM546940
LOCUS
DEFINITION
AGENCOURT_6491295 NIH_MGC_125 Homo sapiens cdna clone IMAGE:5723711
5', mRNA sequence.
BM546940
ACCESSION
BM546940
VERSION
BM546940.1 GI:18780310
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1060)
NIH-MGC http://img.ncbi.nlm.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM12711 row: m column: 24
High quality sequence stop: 737.

FEATURES

source

Location/Qualifiers

1. 1060
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5723711"
/lab_host="DH10B"
/clone.lib="NIH_MGC_125"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORTe;
Site 1: EcorV (destroyed); Site 2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcorV site is oligo-dT primed and directionally cloned
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

ORIGIN

Query Match 45.0%; Score 896.8; DB 12; Length 1060;
Best Local Similarity 94.7%; Pred. No. 2,9e-241;
Matches 973; Conservative 0; Mismatches 37; Indels 17; Gaps 4;

```

QY 488 TTCCCAATCTTATCTGCTGCGAGATGCTCTCAACAGAGCTGATACAGAGA 547
DB 7 TTCCCAATCTTATCTGCTGCGAGATGCTCTCAACAGAGCTGATACAGAGA 66
QY 548 ATGGAGATGGTATGTTAAATGCGAGTATACCTGTCACAAAGCTGACTTATCCCG 607
DB 67 ATGGAGATGGTATGTTAAATGCGAGTATACCTGTCACAAAGCTGACTTATCCCG 126
QY 608 AGTCTCATTTCTGCGTGTGCTGCTGAGTACAGCTTTGAGAAATTTTGCCTGTGT 667
DB 127 AGTCTCATTTCTGCGTGTGCTGCTGAGTACAGCTTTGAGAAATTTTGCCTGTGT 186
QY 668 TGGACAAATCAGTATATTTCAATGAAAGCAAAAGCCCTCAATGATGTTCTAGTGC 727
DB 187 TGGACAAATCAGTATATTTCAATGAAAGCAAAAGCCCTCAATGATGTTCTAGTGC 246
QY 728 ACTGTTAGTGGGATCTCCCGCTCCGCAACATCCCTATCCGCTTATCATATGAAAGGA 787
DB 247 ACTGTTAGTGGGATCTCCCGCTCCGCAACATCCCTATCCGCTTATCATATGAAAGGA 306
QY 788 TGGACATGCTTTTGAAGAGCTTACAGATTGTGAAAGAAAAAGACCTATATATCTC 847
DB 307 TGGACATGCTTTTGAAGAGCTTACAGATTGTGAAAGAAAAAGACCTATATATCTC 366
QY 848 CAAACTTCAATTTTGGGCAACTCTGAGCATGAGAGAAAGTTTAAAGACCAAGCTG 907
DB 367 CAAACTTCAATTTTGGGCAACTCTGAGCATGAGAGAAAGTTTAAAGACCAAGCTG 426
QY 908 GAGCATCAGGGCCAAAGAGCAAACTCAGCTGCTGCACTTGGAGAACCAATGAACTG 967
DB 427 GAGCATCAGGGCCAAAGAGCAAACTCAGCTGCTGCACTTGGAGAACCAATGAACTG 486
QY 968 TCCCTGCTCTCAGAGGGTGGACAGAAACGAGAGCCCTCAGTCCACCTGTGCGG 1027
DB 487 TCCCTGCTCTCAGAGGGTGGACAGAAACGAGAGCCCTCAGTCCACCTGTGCGG 546
QY 1028 ACTGCTACTACAGAGGACAGAGCAAAAGCCCTGATCCCGCAGCTGTCAGAGG 1087
DB 547 ACTGCTACTACAGAGGACAGAGCAAAAGCCCTGATCCCGCAGCTGTCAGAGG 597
QY 1088 TGGCCAGGTGTCAGCCCTGCTGTTAGAGAGACAGCCCGTGGTAAAGGGGCTCACTGAGGCG 1147
DB 598 TGGCCAGGTGTCAGCCCTGCTGTTAGAGAGACAGCCCGTGGTAAAGGGGCTCACTGAGGCG 657
QY 1148 TGCACCTGTCCGAGAGGCTGGAAGACAGCAATTAAGCTCAAGGCTTCTCTCTG 1207
DB 658 TGCACCTGTCCGAGAGGCTGGAAGACAGCAATTAAGCTCAAGGCTTCTCTCTG 717
QY 1208 ATATCAATATGATTTATATTCAGCAGATGAGCATCTTCAAGGCTTCTCTCAT 1267

```

```

DB 718 ATATCAATATGATTTATATTCAGCAGATGAGCATCTTCAAGGCTTCTCTCAT 777
QY 1268 CAGAAAGTGTGTTGAATATCTACAAACCTTCCACTACTGTGAGAGGACCAAGCTTAT 1327
DB 778 CAGAAAGTGTGTTGAATATCTACAAACCTTCCACTACTGTGAGAGGACCAAGCTTAT 837
QY 1328 GCCAGTTCTCCCTGTTCAGAAACATATGAGAGGCAAGTCTCCGAAACCAAGCTTATAG 1387
DB 838 GCCAGTTCTCCCTGTTCAGAAACATATGAGAGGCAAGTCTCCGAAACCAAGCTTATAG 896
QY 1388 AGGAAGCCAGCATTTCCCAAGAGCTGAGACCGCCA-GGCTTCAGAGCCAGAGCAAG 1446
DB 897 AGGAAGCCAGCATTTCCCAAGAGCTGAGACCGCCAAGGCTTCAGAGCCAGAGCAAG 956
QY 1447 CG-----ATTGATTCGTGCTCAGAACAGAGAGAGTGGACCGCCGAGAGTCCCTTTTA 1500
DB 957 CGGATTGATTCGNTGAGAACCCAGAGAGTGGAGCAGCCGCGCAGANGTCCCTTTT 1016
QY 1501 TCTCCAC 1507
DB 1017 TATCCTC 1023

```

RESULT 2
B0945892
LOCUS
DEFINITION
B0945892
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AGENCOURT_8926314 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6463045
B0945892
EST.
B0945892.1 GI:22361370
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 883)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNLN)
Clone distribution by: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
Plate: LNCM2653 row: o column: 14
High quality sequence stop: 672.

FEATURES

source

Location/Qualifiers

1. 883
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6463045"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone.lib="NIH_MGC_101"
/note="Organ: lung; Vector: pOT87; Site 1: EcorI; Site 2:
into EcorI/XhoI sites using the following 5' adaptor:
GGCAAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Query Match 41.8%; Score 834.6; DB 13; Length 883;
Best Local Similarity 98.5%; Pred. No. 9.4e-224;
Matches 874; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

```

QY 794 TGTCTTAGATGAAGCTTACAGATTGTGTAAGAAAAAGAACCTACTATATCTCCAACT 853
DB 1 TGTCTTAGATGAAGCTTACAGATTGTGTAAGAAAAAGAACCTACTATATCTCCAA--C 57
QY 854 TCATATTTTCTGGGCAAACTCTGAGTATGAGAAGATTAAAGAACCAAGCTGGAGCAT 913
DB 58 TTTCAATTTTCTGGGCAAACTCTGAGTATGAGAAGATTAAAGAACCAAGCTGGAGCAT 117
QY 914 CAGGGCCAAAGAGCAAACTCAAGCTGTGCACTCTGAGAAAGCAAAATGAATCTGTCTTG 973
DB 118 CAGGGCCAAAGAGCAAACTCAAGCTGTGCACTCTGAGAAAGCAAAATGAATCTGTCTTG 177
QY 974 CTGTCTCAGAGGGGTGACAGAAAAAGAGAGCCCTGAGTCCAGCTGTGAGGAGCTGTG 1033
DB 178 CTGTCTCAGAGGGGTGACAGAAAAAGAGAGCCCTGAGTCCAGCTGTGAGGAGCTGTG 227
QY 1034 CTACCTCAGAGGAGAGAGCAAGAAAGCCCTGAGTCCCTGAGCTGAGGAGCTGAGCA 1093
DB 238 CTACCTCAGAGGAGAGAGCAAGAAAGCCCTGAGTCCCTGAGCTGAGGAGCTGAGCA 297
QY 1094 GCGTGCAGCCGCTGCTGTAGAGAGACGCCGCTGTGATCAGGGCTGATGGGCTGCACC 1153
DB 298 GCGTGCAGCCGCTGCTGTAGAGAGACGCCGCTGTGATCAGGGCTGATGGGCTGCACC 357
QY 1154 TGTCCGAGACAGGCTGGAGAGAGCAATAAGCTCAAGCGTTCTCTCTGTGATATCA 1213
DB 358 TGTCCGAGACAGGCTGGAGAGAGCAATAAGCTCAAGCGTTCTCTCTGTGATATCA 417
QY 1214 AATCAGTTTATATTTAGCAGAGATGAGAGATCTTACATAGGCTTCTCTCATCAGAA 1273
DB 418 AATCAGTTTATATTTAGCAGAGATGAGAGATCTTACATAGGCTTCTCTCATCAGAA 477
QY 1274 ATGCTTTGAATATCTACAACTTTTCACTACTGTGATGGAGCAACAACTATGCCAGT 1333
DB 478 ATGCTTTGAATATCTACAACTTTTCACTACTGTGATGGAGCAACAACTATGCCAGT 537
QY 1334 TTTCTCCCTGTTTTCAGAACTATCCGAGAGACCTCCGAAACCAAGTCTGATTAAGAGAA 1393
DB 538 TTTCTCCCTGTTTTCAGAACTATCCGAGAGACCTCCGAAACCAAGTCTGATTAAGAGAA 597
QY 1394 CCAGCATCTCCCAAGAACTGTCAGACCGCCAGGCTTTCAGACAGCAGAGAGAGATTGC 1453
DB 598 CCAGCATCTCCCAAGAACTGTCAGACCGCCAGGCTTTCAGACAGCAGAGAGAGATTGC 657
QY 1454 ATTGCTGACAAACAGAGAGAGTGGACCGCCCAAGAGTCTCTTTATCTCCACTGCATC 1513
DB 658 ATTGCTGACAAACAGAGAGAGTGGACCGCCCAAGAGTCTCTTTATCTCCACTGCATC 717
QY 1514 GAAGTGGAGAGCTGGAGAGACAAATTAACACACAGCTTCTTTTGGCTTTCCACAGCC 1573
DB 718 GAAGTGGAGAGCTGGAGAGACAAATTAACACACAGCTTCTTTTGGCTTTCCACAGCC 777
QY 1574 AGCAGCAGCTCAAGAGTCTGCTGGAGCTTTAAGAGCTGGCACTCGGATATCTTTG 1633
DB 778 AGCAGCAGCTCAAGAGTCTGCTGGAGCTTTAAGAGCTGGCACTCGGATATCTTTG 836
QY 1634 CCCCCCAGACCTTACCTCTTCCCTGACACAGACAGTGTATTTTGGC 1680
DB 837 CCCCCCAGACCTTACCTCTTCCCTGACACAGACAGTGTATTTTGGC 883

```

```

REFERENCE 1 (bases 1 to 881)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LCM2648 row: h column: 21
High quality sequence step: 676.
Location/Qualifiers
1. 881
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6460964"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 101"
/note="Organ: lung; Vector: pOTB1; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming, directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

```

ORIGIN

Query Match 41.5%; Score 827.8; DB 13; Length 881;
Best Local Similarity 99.4%; Pred. No. 7,8e-222;
Matches 862; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

```

QY 794 TGTCTTAGATGAAGCTTACAGATTGTGTAAGAAAAAGAACCTACTATATCTCCAACT 853
DB 1 TGTCTTAGATGAAGCTTACAGATTGTGTAAGAAAAAGAACCTACTATATCTCCAACT 60
QY 854 TCATATTTTCTGGGCAAACTCTGAGTATGAGAAGATTAAAGAACCAAGCTGGAGCAT 913
DB 61 TCATATTTTCTGGGCAAACTCTGAGTATGAGAAGATTAAAGAACCAAGCTGGAGCAT 120
QY 914 CAGGGCCAAAGAGCAAACTCAAGCTGTGCACTCTGAGAAAGCAAAATGAATCTGTCTTG 973
DB 121 CAGGGCCAAAGAGCAAACTCAAGCTGTGCACTCTGAGAAAGCAAAATGAATCTGTCTTG 180
QY 974 CTGTCTCAGAGGGGTGACAGAAAAAGAGAGCGCTTCACTGATCCCTGTGCGGACTCTG 1033
DB 181 CTGTCTCAGAGGGGTGACAGAAAAAGAGAGCGCTTCACTGATCCCTGTGCGGACTCTG 240
QY 1034 CTACCTCAGAGGAGAGAGCAAAAGCCGCTGATCCCGCAGCGTCCAGGAGTGGCCA 1093
DB 241 CTACCTCAGAGGAGAGAGCAAAAGCCGCTGATCCCGCAGCGTCCAGGAGTGGCCA 300
QY 1094 GCGTGCAGCCGCTGCTGTAGAGAGACGCCGCTGTATCAGAGGCTCAAGTGGGCTGCACC 1153
DB 301 GCGTGCAGCCGCTGCTGTAGAGAGACGCCGCTGTATCAGAGGCTCAAGTGGGCTGCACC 360
QY 1154 TGTCCGAGACAGGCTGGAGAGAGCAATAAGCTCAAGCGTTCTCTCTGTGATATCA 1213
DB 361 TGTCCGAGACAGGCTGGAGAGAGCAATAAGCTCAAGCGTTCTCTCTGTGATATCA 420
QY 1214 AATCAGTTTATATTTAGCAGAGATGAGAGATCTTACATAGGCTTCTCTCATCAGAA 1273
DB 421 AATCAGTTTATATTTAGCAGAGATGAGAGATCTTACATAGGCTTCTCTCATCAGAA 480
QY 1274 ATGCTTTGAATATCTACAACTTTTCACTACTGTGATGGAGCAACAACTATGCCAGT 1333
DB 481 ATGCTTTGAATATCTACAACTTTTCACTACTGTGATGGAGCAACAACTATGCCAGT 540

```

RESULT 3
BQ930140 881 bp mRNA linear EST 20-AUG-2002
LOCUS BQ930140
DEFINITION AGENCOURT 8923732 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6460964
5', mRNA sequence.
ACCESSION BQ930140
VERSION BQ930140.1 GI:22345171
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euleleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

QY 1334 TCTCCCTGTTCAGGAATCTATGAGCAGACTCCCGAACCAGTCTGATAGAGGAAG 1393
Db 541 TCTCCCTGTTCAGGAATCTATGAGCAGACTCCCGAACCAGTCTGATAGAGGAAG 600
QY 1394 CAGACATCCCAAGAAAGCTGAGACCGCCAGGCTTCAAGCAGCCAGCAAGCAATTC 1453
Db 601 CCAGATCCCAAGAAAGCTGAGACCGCCAGGCTTCAAGCAGCCAGCAAGCAATTC 660
QY 1454 ATTGGGTCAAGAACCAAGCAGCTGAGACCGCCAGGCTTCAAGCAGCCAGCAATTC 1513
Db 661 ATTGGGTCAAGAACCAAGCAGCTGAGACCGCCAGGCTTCAAGCAGCCAGCAATTC 720
QY 1514 GAAGTGGAGGCTGAGAGCAATTAACACACAGCTTCC-TTTTGGGCTTTCCAGCAGC 1572
Db 721 GAAGTGGAGGCTGAGAGCAATTAACACACAGCTTCC-TTTTGGGCTTTCCAGCAGC 780
QY 1573 CAGGAGCAGCTCAAGAGTCTGCTGGCTT-GGGCTTAAAGGCTGAGCACTCGAATATCTT 1631
Db 781 CAGGAGCAGCTCAAGAGTCTGCTGGCTTAAAGGCTGAGCACTCGAATATCTT 840
QY 1632 GG-CCCCCAGACCTTACCCCTTCCC 1657
Db 841 GGCCCCCAGACCTTACCCCTTCCC 867

RESULT 4
B0933499 859 bp mRNA linear EST 21-ANG-2002
LOCUS AGENCOURT_8732489 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6455595
DEFINITION 5', mRNA sequence.
ACCESSION B0933499
VERSION B0933499.1 GI:22348882
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM2634 row: 1 column: 04
High quality sequence stop: 727.
Location/Qualifiers
1. 859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6455595"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_101"
/note="Organ: lung; Vector: POTB1; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

ORIGIN
Query Match 40.6%; Score 809.6; DB 13; Length 859;

Beet Local Similarity 98.5%; Pred. No. 1,1e-216;
Matches 827; Conservative 0; Mismatches 11; Indels 2; Gaps 1;
QY 794 TGTCTTAAATGATAGCTTACAGATTGTGAAAGAAAGAAAGCTATATCTTCAACT 853
Db 1 TGTCTTAAATGATAGCTTACAGATTGTGAAAGAAAGAAAGCTATATCTTCAACT 60
QY 854 TCAATTTCTGGGCCCACTCCCTGACATTTGAGAAAGATTAAAGACAGACTGAGCAT 913
Db 61 TCAATTTCTGGGCCCACTCCCTGACATTTGAGAAAGATTAAAGACAGACTGAGCAT 120
QY 914 CAGGGCCAAAGAGCAAACTCAAGCTGTGACCTGAGAAAGCAAAATGAACCTGTCCCTG 973
Db 121 CAGGGCCAAAGAGCAAACTCAAGCTGTGACCTGAGAAAGCAAAATGAACCTGTCCCTG 180
QY 974 CTGTCTCAGAGGGTGTGACAGAAAAGGAGACGCCCTCAGTCCACCTGTGCGACTGTG 1033
Db 181 CTGTCTCAGAGGGTGTGACAGAAAAGGAGACGCCCTCAGTCCACCTGTGCGACTGTG 240
QY 1034 CTACTCTCAGAGGCGAGCAGACAAAGGCCGTGATCCGCGAGGCGCGAGGCGCGAGGCGCA 1093
Db 241 CTACTCTCAGAGGCGAGCAGACAAAGGCCGTGATCCGCGAGGCGCGAGGCGCGAGGCGCA 300
QY 1094 GCGTGACAGCGGTGCTGTGTAAGAGACAGCCCGCTGTAGAGGCTGACAGTGGCTGAC 1153
Db 301 GCGTGACAGCGGTGCTGTGTAAGAGACAGCCCGCTGTGAGAGAGCGCTGTGACAGGCTGAC 360
QY 1154 TGTGCCAGACAGGCTGTGAGAGCAGCAATTAAGCTCAAGCGTCTCTCTGTGATATCA 1213
Db 361 TGTGCCAGACAGGCTGTGAGAGCAGCAATTAAGCTCAAGCGTCTCTCTGTGATATCA 420
QY 1214 AATCAGTTTCATATTCAAGCAGACATGGAGAGATCTTACATGCTTCTCTCATAGAA 1273
Db 421 AATCAGTTTCATATTCAAGCAGACATGGAGAGATCTTACATGCTTCTCTCATAGAA 480
QY 1274 ATGCTTTGGAAATCTACAAACCTTCACTGATCTGATGAGAACCAAGAGCTATGCACT 1333
Db 481 ATGCTTTGGAAATCTACAAACCTTCACTGATCTGATGAGAACCAAGAGCTATGCACT 540
QY 1334 TCTCCCTGTTCAGGAATCTATGAGCAGACTCCCGAACCAGTCTGATAGAGGAAG 1393
Db 541 TCTCCCTGTTCAGGAATCTATGAGCAGACTCCCGAACCAGTCTGATAGAGGAAG 600
QY 1394 CAGACATCCCAAGAAAGCTGAGACCGCCAGGCTTCAAGCAGCCAGCAAGCAATTC 1453
Db 601 CAGACATCCCAAGAAAGCTGAGACCGCCAGGCTTCAAGCAGCCAGCAAGCAATTC 660
QY 1454 ATTGGGTCAAGAACCAAGCAGCTGAGACCGCCAGGCTTCAAGCAGCCAGCAATTC 1513
Db 661 ATTGGGTCAAGAACCAAGCAGCTGAGACCGCCAGGCTTCAAGCAGCCAGCAATTC 720
QY 1514 GAAGTGGAGGCTGAGAGCAATTAACACACAGCTTCC-TTTTGGGCTTTCCAGCAGC 1573
Db 721 GAAGTGGAGGCTGAGAGCAATTAACACACAGCTTCC-TTTTGGGCTTTCCAGCAGC 780
QY 1574 AGCAGACCTCAAGAGTCTGTGCGC--TGGGCTTAAAGGCTGAGCACTCGAATATCTT 1631
Db 781 AGCAGACCTCAAGAGTCTGTGCGCCTTAAAGGCTGAGNNCACTCGAATATCTT 840

RESULT 5
B0721265 898 bp mRNA linear EST 16-JUL-2002
LOCUS AGENCOURT_8294508 Lupski symptomatic_trunk Homo sapiens cDNA clone
DEFINITION IMAGE:6194455 5', mRNA sequence.
ACCESSION B0721265
VERSION B0721265.1 GI:21860162
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Euteleostomi; Primates; Catarrhini; Homiidae; Homo.
COMMENT (bases 1 to 898)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM13599 row: 1 column: 08
 High quality sequence stop: 669.
 Location/Qualifiers

FEATURES

source

1. 898
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6194455"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /clone_1lb="Lupski, sympathetic trunk"
 /note="Vector: pCMV-Sport6 (Life Technologies); Site_1:
 NotI; Site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCACCGCGTCG-3' and
 5'-GACTAGTTTCAGATCGGAGCGCGCCCT(15)-3'. Size selected >
 1 kb for average insert length 1.9 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine); available through Life
 Technologies."

ORIGIN

Query Match 40.1%; Score 799.8; DB 13; Length 898;

Best Local Similarity 96.8%; Pred. No. 6,4e-214; Indels 6; Gaps 4;
 Matches 859; Conservative 0; Mismatches 22;

QY 8 ATGAGATGATGGAACTCAATGTTGTACTGAGAGGTGGTGGCTGCTGGAAGTGA 67
 DB 12 ATGAGATGATGGAACTCAATGTTGTACTGAGAGGTGGTGGCTGCTGGAAGTGA 71
 QY 68 CGGAAAAAGTGTCTTAATTGATAGCGGCCATTGTGGAATCAATATCCCATTT 127
 DB 72 CGGAAAAAGTGTCTTAATTGATAGCGGCCATTGTGGAATCAATATCCCATTT 131
 QY 128 TGAAGCCATTATATCACTGCTCCCAAGCTTATGAAGCGAAGTTGCAACAGACAA 187
 DB 132 TGAAGCCATTATATCACTGCTCCCAAGCTTATGAAGCGAAGTTGCAACAGACAA 191
 QY 188 TGTAAATTAAGAGCTCATCCAGATTCAAGCAACATAGATTGACATTGTCAGTC 247
 DB 192 TGTAAATTAAGAGCTCATCCAGATTCAAGCAACATAGATTGACATTGTCAGTC 251
 QY 248 AGAAGTTGATGTTAGATCAAGCTCCCAAGATGTGCTCTCTCTTCAAGCTGT 307
 DB 252 AGAAGTTGATGTTAGATCAAGCTCCCAAGATGTGCTCTCTCTTCAAGCTGT 311
 QY 308 TTCTCACTGATCTTGGGTTAACTGAGAAAGAGCTTCACTCTTCACTCTTGCAG 367
 DB 312 TTCTCACTGATCTTGGGTTAACTGAGAAAGAGCTTCACTCTTCACTCTTGCAG 371
 QY 368 GTGGGTTTGTGAGTTCTCTGTTGTTCCCTGGCCTCTGTGAAGAAATCACTCTAG 427
 DB 372 GTGGGTTTGTGAGTTCTCTGTTGTTCCCTGGCCTCTGTGAAGAAATCACTCTAG 431
 QY 428 TCCCTACCTGATTTCTCAGCTTGTCTTACCTGTGTCACATTTGGGCCAACCCGAATTC 487
 DB 432 TCCCTACCTGATTTCTCAGCTTGTCTTACCTGTGTCACATTTGGGCCAACCCGAATTC 491

QY 488 TTCCCAATCTTATCTTGGCTGGCAGAGATGTCCTCAACAGAGCTGATACAGAGA 547
 DB 492 TTCCCAATCTTATCTTGGCTGGCAGAGATGTCCTCAACAGAGCTGATACAGAGA 551
 QY 548 ATGGAGTTGGTATATGTTAAATGCGAGCTATACCTGTCCAAAGCCTGATTTATCCCG 607
 DB 552 ATGGAGTTGGTATATGTTAAATGCGAGCTATACCTGTCCAAAGCCTGATTTATCCCG 611
 QY 608 AGCTCATTTTCCCTGCTGCTGCTGTGAATGACAGCTTTTGTGAATAATTTCCCGTGT 667
 DB 612 AGCTCATTTTCCCTGCTGCTGCTGTGAATGACAGCTTTTGTGAAG-AAATTTCCCGTGT 670
 QY 668 TGAACAATCAGATATTCATTGATGAGAAAGCAAAAGCTCCAAATGATGTTCTAGTGC 727
 DB 671 TGAACAATCAGATATTCATTGATGAGAAAGCAAAAGCTCCAAATGATGTTCTAGTGC 730
 QY 728 ACTGTTTACCTGGATCTCCCGCTCCGCAACATGCTATCGCTACATCATGAGAAGA 787
 DB 731 ACTGTTTACCTGGATCTCCCGCT-CGGACCATGCTATCGCTACATCATGAGAAGA 789
 QY 788 TGAACATGTC-TTAAATGAGCTTACAGATTTGTGAAGAAAGAAAGACCTACTATATCT 846
 DB 790 TGAACATGTC-TTAAATGAGCTTACAGATTTGTGAAGAAAGAAAGACCTACTATATCT 849
 QY 847 CCAACTT---CAATTTCTGGGCAACTCTGGAATGAGAAAGA 890
 DB 850 CCAAACTTCAATTTCTTGGGCAACTCTGGAATGAGAAAGA 896

RESULT 6
 B0951695
 LOCUS
 DEFINITION AGENCOURT_9947130 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6461668
 5', mRNA sequence.

ACCESSION B0951695
 VERSION B0951695.1 GI:22367173
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.
 TITLE NIH-MGC <http://mgc.nci.nih.gov/>.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM2650 row: f column: 05
 High quality sequence stop: 617.
 Location/Qualifiers

FEATURES

source

1. 974
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6461668"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_101"
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
 XhoI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACAGC(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."


```

Db      362 CCGGACGTAAGAGAGAGATTAAGAACGACGATCGAGGCGCAAGGCAAACT 421
QY      933 CAAGCTGCTGACCTGAGAGAGCCAAATGAACCTGCTGCTGCTGCTGAGAGGCTGACA 992
Db      422 CAAGCTGCTGACCTGAGAGAGCCAAATGAACCTGCTGCTGCTGCTGAGAGGCTGACA 481
QY      993 GAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1052
Db      482 GAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
QY      1053 ACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1112
Db      542 ACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
QY      1113 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1172
Db      602 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661
QY      1173 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1231
Db      662 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721
QY      1232 CCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1287
Db      722 CCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781
QY      1288 TACAAAGCTTCCACTACTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1342
Db      782 TACAAAGCTTCCACTACTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
QY      1343 TTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1402
Db      842 TTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
QY      1403 C 1403
Db      902 C 902

```

```

RESULT 8
BI821804 836 bp mRNA linear EST 04-OCT-2001
LOCUS 60305883F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176724 5'
DEFINITION mRNA sequence.
ACCESSION BI821804
VERSION BI821804.1 GI:15933354
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 836)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: C9apbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Inceye Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1440 row: f column: 21
High quality sequence stop: 805.
Location/Qualifiers
1. 836
FEATURES
Source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5176724"

```

```

/lab host="DH10B"
/notes="Organ: pooled brain, lung, testis; Vector:
pcmv-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 34.9%; Score 695.6; DB 12; Length 836;
Best Local Similarity 96.1%; Pred. No. 1,7e-184;
Matches 809; Conservative 0; Mismatches 24; Indels 9; Gaps 9;

147 CTGCTCAAGCTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 206
1 CTGCTCAAGCTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
207 CCAGAGATTCAGCGAAACATAGAGTTGACATTTAGTCAGTCAGAGAGTTGTAGTTACGA 266
61 CCAGAGATTCAGCGAAACATAGAGTTGACATTTAGTCAGTCAGAGAGTTGTAGTTACGA 119
267 TCAAGCTCCCAAGATGTTGCTCTCTCTTGAAGAGTGTCTGACTGTCTGACTGTCTGAGG 326
120 TCAAGCTCCCAAGATG-TGCCCTCTCTCTCTGAGAGTGTCTGACTGTCTGAGG 178
327 TAAAGTGAAGAGAGTTCAACTCTGTCACCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 386
179 TAAAGTGAAGAGAGTTCAACTCTGTCACCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAG 238
387 TCGTGTGTTCCCTGAGCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 446
239 TCGTGTGTTCCCTGAGCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 298
447 GCTTGTGTTCCCTGAGCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 506
299 GCTTGTGTTCCCTGAGCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 358
507 CTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 566
359 CTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 418
567 AATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626
419 AATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478
627 GCTTGTGTTCCCTGAGCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 686
479 GCTTGTGTTCCCTGAGCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 538
687 CATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 746
539 CATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597
747 CCGCTCCGACACCATGCTATGCTCTATCATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 806
598 CCGCTCCG-CACCATGCTATGCTCTATCATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 864
807 AGCTTACAGATTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 924
656 AGCTTACAGATTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 984
716 GGCACAGCTCTGAGCTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
865 GGCACAGCTCTGAGCTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 924
925 AGCAAACTCAAGCTGTGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 984
775 AGCTTAATCAAGCTGTGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 833

```


RESULT	10
LOCUS	BE897795
DEFINITION	BE897795 920 bp mRNA linear EST 20-OCT-2000
ACCESSION	601438457F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3923247 5,'
VERSION	BE897795
KEYWORDS	BE897795.1 GI:10363618
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (Baes 1 to 920)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: ATCC/DCTD/DTP

CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM9759 row: b column: 16
 High quality sequence stop: 678.

FEATURES

Source

Location/Qualifiers

1..920
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3923247"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: oligo dr.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match 32.1%; Score 641.2; DB 10; Length 920;
 Best Local Similarity 95.0%; Pred. No. 4,1e-169;
 Matches 709; Conservative 0; Mismatches 23; Indels 14; Gaps 4;

699 AAAAACTCCATGATGATGTTCTTCTAGTGCACCTGTTAGCTGGATCTCCCGTCCGAC 758
 5 AAGGCTCCCGGTGATGTTCTTCTAGTGCACCTGTTAGCTGGATCTCCCGTCCGAC 64
 759 CATGCTATGCTGCTATCATGATGAAGGATGACATGCTTTAGATGAAGTCAAGATT 818
 65 CATGCTATGCTGCTATCATGATGAAGGATGACATGCTTTAGATGAAGTCAAGATT 124
 819 TGTGAAGAAAAAGAACTCTATATCTCCAACTTCAATTTCTGGGCCAACTCTCGA 878
 125 TGTGAAGAAAAAGAACTCTATATCTCCAACTTCAATTTCTGGGCCAACTCTCGA 184
 879 CTATGAAGAAAGATTAAAGCAAGATGAGATCAAGGATCAAGGATCAAGGATCAAG 938
 185 CTATGAAGAAAGATTAAAGCAAGATGAGATCAAGGATCAAGGATCAAGGATCAAG 244
 939 GCTGCACTGGAAGCAAGCAATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 998
 245 GCTGCACTGGAAGCAAGCAATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 304
 999 CGAGAGCCCTCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1058
 305 CGAGAGCCCTCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364
 1059 GCGCTGATGATCCGCGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAG 1118
 365 GCGCTGATGATCCGCGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAG 415
 1119 CAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1178
 416 CAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 475
 1179 CAATTAAGCTCAAGCGTCTCTCTCTGATATCAATTAAGCTCAATTAAGCTCAAT 1238
 476 CAATTAAGCTCAAGCGTCTCTCTCTGATATCAATTAAGCTCAATTAAGCTCAAT 535
 1239 GCGAGATCTCTTAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1298
 536 GCGAGATCTCTTAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 595
 1299 CACTACTCTGATGAGGAGCAAGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1358
 596 CACTACTCTGATGAGGAGCAAGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 654
 1359 GCAGACTCCCGAAACCAAGCTCTGATTAAGAGAGCAAGCAAGCAAGCAAGCTGCA 1418

Db 655 GCAGACTCCCGAAC--AGTCTGATGAGAGAGCAAGCAGCATCCCAAGAGGTGAGAC 711
 QY 1419 CGCCAGCCCTTCAGACGACGAGCA 1444
 Db 712 CG-CAGGCGCTTCAGAGCCGAGCA 736

RESULT 11

LOCUS

DEFINITION

BI816954 682 bp mRNA linear EST 10-OCT-2001
 imageqc.10.2000/s1z410bdf141.x1 Soares.NPBC Homo sapiens CDNA

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The I.M.A.G.E. Consortium quality control effort: clone
 resequencing for verification
 Unpublished (2001)
 Other ESTs: BC058779

CONTACT

PRANGE CK

THE I.M.A.G.E. CONSORTIUM

LAWRENCE LIVERMORE NATIONAL LABORATORY

LIVERMORE, CA, USA

EMAIL: help@image.llnl.gov

This read has been verified (found to hit its original self in the
 correct orientation), as part of the I.M.A.G.E. Consortium quality
 control effort. High quality sequence is defined as having 100 or
 more base pairs with a phred quality value of 20 or greater, where
 a sliding window of 4 base pairs with a phred quality value of 15
 or greater marks the beginning and end of the sequence. For
 information on obtaining this clone, please contact
 info@image.llnl.gov effort.

PLATE: LLM9388

ROW: K

COLUMN: 7

SEQ PRIMER: -21m13

HIGH QUALITY SEQUENCE STOP: 682.

LOCATION/QUALIFIERS

1..682

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4140798"

/tissue_type="lymphocyte"

/lab_host="DH10B (phage-resistant)"

/note="Organ: Blood; Vector: pTT3D-Pac; Site 1: NotI;
 Site 2: EcoRI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5'
 TGTACCAATCTGAAGTGGAGGCGCGCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified pTT3 vector. Library
 is normalized; constructed in the laboratory of M. Bento
 Soares (University of Iowa)."

ORIGIN

Query Match 31.7%; Score 633; DB 12; Length 682;
 Best Local Similarity 99.1%; Pred. No. 6.9e-167;
 Matches 647; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 169 AGGTGCAACAGCAAGAGTGAATTAAGAGCTCATCAGCATTCAGGAAACATTAAG 228
 Db 652 AAGGTGCAACAGCAAGAGTGAATTAAGAGCTCATCAGCATTCAGGAAACATTAAG 594
 QY 229 GTTGACATTGATTGACAGTCAAGAGTGTGATTACGATCAAGCTCCCAAGATGTTGCC 288
 Db 593 GTTGACATTGATTGACAGTCAAGAGTGTGATTACGATCAAGCTCCCAAGATGTTGCC 534

```

QY 289 TCTCTCTGAGAGCTGTTTCTCACTGTAATTCTGGTAACGTGAGAGAGCTTCAAC 348
Db 533 TCTCTCTCTGAGAGCTGTTTCTCACTGTAATTCTGGTAACGTGAGAGAGCTTCAAC 474
QY 349 TCTGTTCACTGCTTGGAGGTGAGTGTGCTGAGTGTCTGCTGTTTCTCTGCTGCTGCT 408
Db 473 TCTGTTCACTGCTTGGAGGTGAGTGTGCTGAGTGTCTGCTGTTTCTCTGCTGCTGCT 414
QY 409 GAGAGAAATCACTCACTGAGTCCCTTACCTGATTTTCTGACCTTGTCTTACTGTTGCAAC 468
Db 413 GAGAGAAATCACTCACTGAGTCCCTTACCTGATTTTCTGACCTTGTCTTACTGTTGCAAC 354
QY 469 ATTGGGCAACCCGGAATTTCTTCCCACTTTTATCTTGGCTGACGAGAGATGCTCAAC 528
Db 353 ATTGGGCAACCCGGAATTTCTTCCCACTTTTATCTTGGCTGACGAGAGATGCTCAAC 294
QY 529 AAGAGCTGATACAGAGAGAGAGATGATGTTATGTTTAAATGCCAGCTATACCTGTCA 588
Db 293 AAGAGCTGATACAGAGAGAGAGATGATGTTATGTTTAAATGCCAGCTATACCTGTCA 234
QY 589 AAGCTGACTTTATCCCGAGTCTCATTTCTGCGCTGCTGCTGAGTATGACGCTTTTGT 648
Db 233 AAGCTGACTTTATCCCGAGTCTCATTTCTGCGCTGCTGCTGAGTATGACGCTTTTGT 174
QY 649 GAGAAATTTTGCCTGCTGCTGCAAAATCACTGATTTTCAATTGAGAAAGCAAAAGCTCC 708
Db 173 GAGAAATTTTGCCTGCTGCTGCAAAATCACTGATTTTCAATTGAGAAAGCAAAAGCTCC 114
QY 709 AATGATGTTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 768
Db 113 AATGATGTTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 54
QY 769 GCTTACATCAAGAGAGAGAGAGATGATGTTTATGATGAGCTTACAGATTTGT 821
Db 53 GCTTACATCAAGAGAGAGAGAGATGATGTTTATGATGAGCTTACAGATTTGT 1

```

```

RESULT 12
LOCUS BK479029 663 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686J2208.r1.686 (synonym: hicc3) Homo sapiens cDNA clone
ACCESSION BKZp686J2208.5, mRNA sequence.
VERSION BK479029
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
JOURNAL EST (Bahr, A., Lauber, J., Mewes, H.W., Weil, B., et al.)
COMMENT Unpublished (2003)
CONTACT: MIPS
MIPS

```

Ingolstaedter Landstr. 1, D-85764 Neuberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de,
 consortium of the German Genome Project.
 No 5' sequence available.
 This clone (DKFZp686J2208) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```

FEATURES
SOURCE
1..663
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686J2208"

```

```

/dev_stage="adult"
/lab_host="DH10B"
/clone_id="686 (synonym: hicc3)"
/notes="Vector: pTIPLEX2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

```

```

Query Match 31.7%; Score 632; DB 13; Length 663;
Best Local Similarity 98.6%; Pred. No. 1,3e-16;
Matches 647; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

```

```

QY 665 GGTGAGCAATCGATGATTTTCAAT--GAGAAAGCAAAAGCTTCAATGATGTTTCT 722
Db 2 GGTGAGCAATCGATGATTTTCAATGAGAAAGCAAAAGCTTCAATGAGATGTTTCT 61
QY 723 AGTGCATGTTTACCTGAGATCTCCGCTCCGCAACATGCGATATGCTTACATATGAA 782
Db 62 AGTGCATGTTTACCTGAGATCTCCGCTCCGCAACATGCGATATGCTTACATATGAA 121
QY 783 GAGATGAGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
Db 122 GAGATGAGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181
QY 843 ATCTCCAACTTCAATTTTCTGGGCAACTCTGCACTATGAGAGAGATTTAGAACCA 902
Db 182 ATCTCCAACTTCAATTTTCTGGGCAACTCTGCACTATGAGAGAGATTTAGAACCA 241
QY 903 GACTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 962
Db 242 GACTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
QY 963 ACTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
Db 302 ACTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
QY 1023 TGGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
Db 362 TGGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
QY 1083 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142
Db 422 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
QY 1143 TGGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
Db 482 TGGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
QY 1203 TCTGATATCAATCACTTTCATATTCAGCCAGCATGCGATGCTTACATGCTTCTC 1262
Db 542 TCTGATATCAATCACTTTCATATTCAGCCAGCATGCGATGCTTACATGCTTCTC 601
QY 1263 CTCATCAGAGATGCTTGGAAATCACTTCCACTTCTGATGAGGAGCA 1318
Db 602 CTCATCAGAGATGCTTGGAAATCACTTCCACTTCTGATGAGGAGCA 657

```

```

RESULT 13
LOCUS B0670989 891 bp mRNA linear EST 15-JUN-2002
DEFINITION AGENCOURT 8039377 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6211589
ACCESSION B0670989
VERSION B0670989.1 GI:21781823
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://nigc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

```



```
QY 394 TTCCTGGGCTCTGTGAAGAAATCACTAGTCCCTACCTGATTTCTCAGCCTTGC 453
DB 123 TTCCTGGGCTCTGTGAAGAAATCACTAGTCCCTACCTGATTTCTCAGCCTTGC 182
QY 454 TTACTGTTGGCCAACTATGGGCAACCCGAAATCTTCCCAATCTTTATCTTGGCTGCAG 513
DB 183 TTACTGTTGGCCAACTATGGGCAACCCGAAATCTTCCCAATCTTTATCTTGGCTGCAG 242
QY 514 CGAAGTGTCTTCAACAAAGAGCTGATAGAGAAATGGAGATTGGTTATGTGTAAATGCC 573
DB 243 CGAAGTGTCTTCAACAAAGAGCTGATAGAGAAATGGAGATTGGTTATGTGTAAATGCC 302
QY 574 AGCTATACCTGTCCAAAGCCTGACTTTATCCCGAGCTCATTTCTCGCTGTGCTGTCG 633
DB 303 AGCAATACCTGTCCAAAGCCTGACTTTATCCCGAGCTCATTTCTCGCTGTGCTGTCG 362
QY 634 AATGACAGCTTTGTGAGAAATTTTGGCGGTGTGGAACAATGATGATTTCAATTGAG 693
DB 363 AATGACAGCTTTGTGAGAAATTTTGGCGGTGTGGAACAATGATGATTTCAATTGAG 422
QY 694 AAAGCAAAAGCCTTCAATGATGATGTTCTGATGCACTGTTAGCTGGATTTCCCGCTGC 753
DB 423 AAAGCAAAAGCCTTCAATGATGATGTTCTGATGCACTGTTAGCTGGATTTCCCGCTGC 482
QY 754 GCCACATTCGCTATTCGCTTATCATCATGAAGAGATGACATGTCTTTAGATGAAGCTTAC 813
DB 483 GCCACATTCGCTATTCGCTTATCATCATGAAGAGATGACATGTCTTTAGATGAAGCTTAC 542
QY 814 AGATTGTGAAGAAAGAAAGACCTACTATATCTTCCAACTGTAATTTTGTGGGCCAACTC 873
DB 543 AGATTGTGAAGAAAGAAAGACCTACTATATCTTCCAACTGTAATTTTGTGGGCCAACTC 602
QY 874 CTGACATATGAAGAAAGATTAAGAACCAAGCTGAGATCAAGAGGCCAAAGACAACTC 933
DB 603 ATGACATATGAAGAAAGATTAAGAACCAAGCTGAGATCAAGAGGCCAAAGACAACTC 662
QY 934 AACCTGTGACCTGTGAAGAGCCAAATGAAACCTGTCCTGCTGTGCAAGGCTGACAG 993
DB 663 AACCTGTGACCTGTGAAGAGCCAAATGAAACCTGTCCTGCTGTGCAAGGCTGACAG 722
QY 994 AAAAGCGAAGCGCCCTCAGTCCACCCCTGTGCCACTC 1031
DB 723 AAAAGCGAAGCGCCCTCAGTCCACCCCTGTGCCACTC 760
```

```
RESULT 15
BI917706/c 660 bp mRNA linear EST 16-OCT-2001
LOCUS 60318366F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5247884 5',
DEFINITION mRNA sequence.
ACCESSION BI917706
VERSION BI917706.1 GI:16181504
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 660)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strassberg, Ph.D.
Email: cgsdbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.liml.gov
Plate: LIML1625 row: k column: 21
High quality sequence stop: 660.
Location/Qualifiers
```

FEATURES

```
source
1..660
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5247884"
/lab_host="DH10B"
/clone_1id="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."
```

```
Query Match 30.8%; Score 615.4; DB 12; Length 660;
Best Local Similarity 98.2%; Pred. No. 6,3e-162;
Matches 657; Conservative 0; Mismatches 1; Indels 11; Gaps 3;
```

```
QY 707 CCAATGATGTTGTTAGTGAAGCACTTTAGCTGGATCTCCGCTCC-GCCACATCGCT 765
DB 660 CCAATGATGTTGTTAGTGAAGCACTTTAGCTGGATCTCCGCTCCGCTCCGCTCCGCT 601
QY 766 ATGCGCTACATCATGAAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAA 825
DB 600 ATGCGCTACATCATGAAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAA 541
QY 826 GAAAGAAAGCTTATATCTCA-AACTTCAATTTTCTGGGCCAACTCTGGACTATGA 884
DB 540 GAAAGAAAGCTTATATCTCA-AACTTCAATTTTCTGGGCCAACTCTGGACTATGA 481
QY 885 GAAAGAAATTAAGAACCAAGCTGAGACATCAGGCGCAAGAGCAAACTCAAGCTGTGCA 944
DB 480 GAAAGAAATTAAGAACCAAGCTGAGACATCAGGCGCAAGAGCAAACTCAAGCTGTGCA 421
QY 945 CCTGGAGAGCCCAATGAACCTGCTCTGCTGTCTCAGAGGCTGACAGAAAGCGAGAC 1004
DB 420 CCTGGAGAGCCCAATGAACCTGCTCTGCTGTCTCAGAGGCTGACAGAAAGCGAGAC 361
QY 1005 GCCCCTCAGTCCACCTGTGCGACTGCTGCTACCTCAGAGCGACAGAGCAAAAGCCCGT 1064
DB 360 GCCCCTCAGTCCACCTGTGCGACTGCTGCTACCTCAGAGCGACAGAGCAAAAGCCCGT 301
QY 1065 GCATCCCGCAGCGTCCCAAGCGTCCAGCGTGCAGCGCTGTGTTAGAGCAAGCC 1124
DB 300 GCATCCCGCAGCGT-----CCAGCGTCAAGCCGTCGCTGTTAGAGCAAGCC 250
QY 1125 GCTGTGTAAGCGGCTCAGTGGCTGCACTGTCGCGAGACAGAGCTGGAAGACCAATTA 1184
DB 249 GCTGTGTAAGCGGCTCAGTGGCTGCACTGTCGCGAGACAGAGCTGGAAGACCAATTA 190
QY 1185 GCTCAAGGCTTCTTCTCTGATATCAATCAATCAATCAATCAATCAATCAATCAATCA 1244
DB 189 GCTCAAGGCTTCTTCTCTGATATCAATCAATCAATCAATCAATCAATCAATCAATCA 130
QY 1245 ATCTTACATGAGCTTCTCTCATAGAAAGATGCTTTGAAATATCAATCAATCAATCA 1304
DB 129 ATCTTACATGAGCTTCTCTCATAGAAAGATGCTTTGAAATATCAATCAATCAATCA 70
QY 1305 TCTGAGTGGACCAACAGATATGCAAGTTCCTCCCTGTTCAGAGACTATCGAGAGAG 1364
DB 69 TCTGAGTGGACCAACAGATATGCAAGTTCCTCCCTGTTCAGAGACTATCGAGAGAG 10
QY 1365 TCCCGAAGC 1373
DB 9 TCCCGAAGC 1
```

Search completed: June 22, 2004, 02:51:54

Tue Jun 22 16:59:31 2004

us-10-029-345a-108_copy_538_2532.rst

Page 13

Job time : 7901.16 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 12:17:10 ; Search time 1159.45 Seconds
(without alignments)
7309.622 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1 atggcccaatgatgatg9.....tggaatcatgatgctcc 1995

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1995	100.0	5450	6	ACC60559
2	1995	100.0	5450	6	ACC60572 Polynucle
3	1991.8	99.8	3059	6	AA515768
4	1991.8	99.8	3496	6	ABK47596
5	1991.8	99.8	3544	5	AA514639
6	1991.8	99.8	3544	6	ABK49402
7	1991.8	99.8	3766	6	ABK14474
8	1991.8	99.8	4790	6	ABN83966
9	1991.8	99.8	5145	5	ABV20833
10	1991.8	99.8	5145	5	ABV21080
11	1991.8	99.8	5145	5	ABV26680
12	1991.8	99.8	5145	5	ABV20978
13	1991.8	99.8	5145	5	ABV21092
14	1991.8	99.8	5145	5	ABV21312
15	1991.8	99.8	5145	5	ABV21316
16	1991.8	99.8	5145	5	ABV26826
17	1991.8	99.8	5145	5	ABV27131
18	1991.8	99.8	5145	5	ABV26923
19	1991.8	99.8	5145	5	ABV27135
20	1991.8	99.8	5145	5	ABV28657
21	1991.8	99.8	5145	5	ABV28227
22	1991.8	99.8	5145	5	ABV26934
23	1990.2	99.8	2732	4	AA094992

24	1990.2	99.8	3104	6	ABN59704	Abn59704 Novel hum
25	1988.6	99.7	2102	7	ADA53105	Ada53105 Human cod
26	1988.6	99.7	2966	4	AAH99685	Aah99685 Human pro
27	1988.6	99.7	5111	6	ACC60521	Acc60521 Polynucle
28	1968.8	98.7	2118	4	AAE30479	Aae30479 Human pro
29	1653.8	82.9	3332	6	ABK48378	Abk48378 CDNA enco
30	1554.8	77.9	2756	6	ACC60560	Acc60560 Polynucle
31	727.6	36.5	749	4	AAH06539	Aah06539 Human pro
32	417.2	20.9	425	5	ABV10726	Abv10726 Human pro
33	415.2	20.8	467	5	ABV40998	Abv40998 Human pro
34	415.2	20.8	467	5	ABV31891	Abv31891 Human pro
35	415.2	20.8	467	5	ABV40849	Abv40849 Human pro
36	411.2	20.6	418	5	ABV11252	Abv11252 Human pro
37	410.2	20.6	461	5	ABV40981	Abv40981 Human pro
38	410.2	20.6	461	5	ABV23398	Abv23398 Human pro
39	410.2	20.6	461	5	ABV41327	Abv41327 Human pro
40	401.2	20.1	408	5	ABV10907	Abv10907 Human pro
41	399.2	20.0	438	5	ABV32059	Abv32059 Human pro
42	395.4	19.8	427	5	ABV10887	Abv10887 Human pro
43	387.8	19.4	2377	7	ACA64956	Aca64956 Human pro
44	387.8	19.4	2377	7	ABX10760	Abx10760 Human pro
45	382	19.1	2453	2	AAT86758	Aat86758 CDNA of t

ALIGNMENTS

RESULT 1	ACC60559	standard; CDNA; 5450 BP.
ID	ACC60559;	
AC	ACC60559;	
DT	19-JUN-2003	(first entry)
XX		
DE	Polynucleotide relating to the invention SEQ ID NO: 108.	
XX		
KW	Gene; ser; antiproliferative; hepatotropic; nephrotropic; antiarthritic;	
KW	antiproliferative; cardiac; cytoskeletal; gene therapy; liver disease;	
KW	proliferative disorder; renal failure; cardiovascular disorder;	
KW	immunological disorder; arthritis; psoriasis; congenital heart defect;	
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200257460-A2.	
XX		
PD	25-JUL-2002.	
XX		
PF	20-DEC-2001; 2001WO-US050459.	
XX		
PR	20-DEC-2000; 2000US-025668P.	
PR	30-MAR-2001; 2001US-028018P.	
PR	01-MAY-2001; 2001US-0287735P.	
PR	05-JUN-2001; 2001US-0295848P.	
PR	25-JUN-2001; 2001US-0300465P.	
XX		
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.	
XX		
PI	Jackson DG, Feder J, Nelson T, Mintler G, Ramanathan C, Lee L;	
PI	Stemets N, Bol D, Schieven G, Fingler J, Todderud CG, Bassolino D;	
PI	Krystek S, Mcatee P, Suchard S, Banas D;	
XX		
DR	WPI; 2002-599721/64.	
XX		
PT	P-PsDB; ABR52381.	
PT	Novel polynucleotides encoding human phosphatase polypeptides useful in	
PT	the prevention or treatment of e.g. proliferative and cardiovascular	
XX	disorders.	
XX	Claim 1; Fig 13; 801p; English.	
XX	The invention relates to a novel isolated nucleic acid comprising a	

CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic, CC
CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC ameliorating the invention is useful for preventing, treating or
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX

Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;

Query Match 100.0%; Score 1995; DB 6; Length 5450;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGCCATGAGATGATTTGAACTCAATTTTACTGAGAGTGTGGCTGTGCTGAA 60
DB 538 ATGGCCATGAGATGATTTGAACTCAATTTTACTGAGAGTGTGGCTGTGCTGAA 537
QY 61 AGTGAACGGAAAAAGTGTCTAATTTAGTACCGGCGATTTTGGAAATCAATACATCC 120
DB 598 AGTGAACGGAAAAAGTGTCTAATTTAGTACCGGCGATTTTGGAAATCAATACATCC 657
QY 121 CACATTTTGGAAAGCCATTAATTAATCACTGTCTCAAGCTTATGAGCGAAGTTGCAACG 180
DB 658 CACATTTTGGAAAGCCATTAATTAATCACTGTCTCAAGCTTATGAGCGAAGTTGCAACG 717
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCAGCATTCAGCGGAAACATTAAGTTGACATGAT 240
DB 718 GACAAAGTGTAAATTAACAGAGCTCATCAGCATTCAGCGGAAACATTAAGTTGACATGAT 777
QY 241 TGCAGTGAAGGTTTGAATTTACATCAAGCTTCCCAAGATGTTCTCTCTCTCA 300
DB 778 TGCAGTGAAGGTTTGAATTTACATCAAGCTTCCCAAGATGTTCTCTCTCTCA 837
QY 301 GACTGTTTCTCACTGACTTCTGGGGTAACTGGGAAAGAGCTTCACTCTGTTCACTG 360
DB 838 GACTGTTTCTCACTGACTTCTGGGGTAACTGGGAAAGAGCTTCACTCTGTTCACTG 897
QY 361 CTGGCAGAGGTTTGTGTAAGTTCTCTGTTGTTTCCCTGAGCTGTGTAAGGAAATCC 420
DB 898 CTGGCAGAGGTTTGTGTAAGTTCTCTGTTGTTTCCCTGAGCTGTGTAAGGAAATCC 957
QY 421 ACTCTAGTCCCTAAGCTGATTTCTCAGCTTGTCTTACCTGTTGCCAATTTGGCCAAAC 480
DB 958 ACTCTAGTCCCTAAGCTGATTTCTCAGCTTGTCTTACCTGTTGCCAATTTGGCCAAAC 1017
QY 481 CGAATTTCTCCCAATTTTATCTTGGCTGCAAGAGATGTCTCAACAAAGAGCTGATA 540
DB 1018 CGAATTTCTCCCAATTTTATCTTGGCTGCAAGAGATGTCTCAACAAAGAGCTGATA 1077
QY 541 CAGCAAGATGAGATGTTATGTTAATGCAAGCTTAATCTGTCCAAAGCTGACTTT 600
DB 1078 CAGCAAGATGAGATGTTATGTTAATGCAAGCTTAATCTGTCCAAAGCTGACTTT 1137
QY 601 ATCCCGAGTCTCATTTCTGCTGTGTGCTGTGATGACAGCTTTTGTGAAATTTTG 1197
DB 1138 ATCCCGAGTCTCATTTCTGCTGTGTGCTGTGATGACAGCTTTTGTGAAATTTTG 1257
QY 661 CCGTGTGTTGACAATCATGATGATTTGATGAAAGCAAAAGCTTCCAAATGATGTT 720
DB 1198 CCGTGTGTTGACAATCATGATGATTTGATGAAAGCAAAAGCTTCCAAATGATGTT 1257
QY 721 CTAGTGAAGTGTGTTAGCTGGGATCTCCGCTCCGCAACCATGCTATGCTCATGATG 780
DB 1258 CTAGTGAAGTGTGTTAGCTGGGATCTCCGCTCCGCAACCATGCTATGCTCATGATG 1317
QY 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGTGAAAGAAAAAGACTTACT 840
```

```
DB 1318 AAGAGATGACATGCTTTAGATGAGCTTATGATTTGTGAAAGAAAAAGCTTACT 1377
QY 841 ATATCTCAAACTTCAATTTTCTGGGCGCACTCTGTGACTATGAGAAAGATTTAAGAC 900
DB 1378 ATATCTCAAACTTCAATTTTCTGGGCGCACTCTGTGACTATGAGAAAGATTTAAGAC 1437
QY 901 CAGATGAGAGCTTACGAGGCGCAAGAGCAAACTCAAGCTGTGCACTGTGAGAAAGCAAT 960
DB 1438 CAGATGAGAGCTTACGAGGCGCAAGAGCAAACTCAAGCTGTGCACTGTGAGAAAGCAAT 1497
QY 961 GAACCTGTCTCCGTGTCTGAGAGGTTGACAAAAAGCAAGCGCCCTCACTCAACC 1020
DB 1498 GAACCTGTCTCCGTGTCTGAGAGGTTGACAAAAAGCAAGCGCCCTCACTCAACC 1020
QY 1021 TGTGCGCACTCTGCTTACCTTCAAGAGGCAAGAGCAAAAGCGCTGTGATCCGCGAGCGTG 1080
DB 1558 TGTGCGCACTCTGCTTACCTTCAAGAGGCAAGAGCAAAAGCGCTGTGATCCGCGAGCGTG 1617
QY 1081 CCAGCGTCCAGCGTGTGAGAGGCGCTGTGATGAGAGCAAGCGCTGTGATGAGAGCGCTG 1140
DB 1618 CCAGCGTCCAGCGTGTGAGAGGCGCTGTGATGAGAGCAAGCGCTGTGATGAGAGCGCTG 1677
QY 1141 AGTGGCTGCACTGTGCTCCGAGCAAGGCTGGAAGA CAGCAATAGCTCAAGGTTCTTC 1200
DB 1678 AGTGGCTGCACTGTGCTCCGAGCAAGGCTGGAAGA CAGCAATAGCTCAAGGTTCTTC 1737
QY 1201 TCTCTGATATCAATCACTTCAATTAATGAGGCAAGGCTGAGCACTTCAATGAGCTTC 1260
DB 1738 TCTCTGATATCAATCACTTCAATTAATGAGGCAAGGCTGAGCACTTCAATGAGCTTC 1797
QY 1261 TCTCTGATATCAATCACTTCAATTAATGAGGCAAGGCTGAGCACTTCAATGAGCTTC 1320
DB 1798 TCTCTGATATCAATCACTTCAATTAATGAGGCAAGGCTGAGCACTTCAATGAGCTTC 1857
QY 1321 AAGCTATGCAAGTCTTCCCTGTTTCAAGAACTATGAGAGCACTCCGAAACAGTCTC 1380
DB 1858 AAGCTATGCAAGTCTTCCCTGTTTCAAGAACTATGAGAGCACTCCGAAACAGTCTC 1917
QY 1381 GATTAAGAGAGAGCCAGATCCCAAGAGCTGAGACCGCCAGGCGCTTCAAGAGCGAG 1440
DB 1918 GATTAAGAGAGAGCCAGATCCCAAGAGCTGAGACCGCCAGGCGCTTCAAGAGCGAG 1977
QY 1441 AGCAAGCATTTGCAATTTGCTGAGAGCAAGAGCTGAGAGCACTCCGAAACAGTCTC 1500
DB 1978 AGCAAGCATTTGCAATTTGCTGAGAGCAAGAGCTGAGAGCACTCCGAAACAGTCTC 2037
QY 1501 TCTCACTGATGAGAGTGGAGAGCTGAGAGCAATTTACCAAGAGCTTCTTTTGGG 1560
DB 2038 TCTCACTGATGAGAGTGGAGAGCTGAGAGCAATTTACCAAGAGCTTCTTTTGGG 2097
QY 1561 CTTTCAAGAGAGAGAGCAAGCTTCAAGAGCTGAGAGCACTTCAAGAGCTTCTTTTGGG 1620
DB 2098 CTTTCAAGAGAGAGAGCAAGCTTCAAGAGCTGAGAGCACTTCAAGAGCTTCTTTTGGG 2157
QY 1621 TCGATATTTTGGGCGCCCAAGAGCTTCAAGAGCTTCAAGAGCTTCTTTTGGG 1680
DB 2158 TCGATATTTTGGGCGCCCAAGAGCTTCAAGAGCTTCAAGAGCTTCTTTTGGG 2217
QY 1681 ACAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCTT 1740
DB 2218 ACAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCTT 2277
QY 1741 GCTTCAAGCTGAGAGCAAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 2337
DB 2278 GCTTCAAGCTGAGAGCAAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 2397
QY 1801 CAGAGAGCAAGAGCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 1860
DB 2338 CAGAGAGCAAGAGCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 2397
QY 1861 AAGCATTTAAAGCAAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 1920
```

Db 2398 AAGCAGCTTTAAACGCGAAGAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTCAAGAAAC 2457
 Oy 1921 AGGTATCCGGGAAGAGCTGGCGGGAATGTGGCAGTCACTAGCTTTTCGGGCGAGCATGGAA 1986
 Db 2458 AGGTCAACCGGAAGAGCTGGCGGGAAGTGGGCGAGTCACTAGCTTTTCGGGCGAGCATGGAA 2517
 Oy 1981 ATCATTTAGAGTCTCC 1995
 Db 2518 ATCATTTAGAGTCTCC 2532
 RESULT 2
 ID ACC60572
 AC ACC60572 standard; cDNA; 5450 BP.
 XX
 AC ACC60572;
 DT 19-JUN-2003 (first entry)
 XX
 DE Polynucleotide relating to the invention SEQ ID NO: 147.
 XX
 KW Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic;
 KW antipneumatic; cardiant; cytostatic; gene therapy; liver disease;
 KW proliferative disorder; renal failure; cardiovascular disorder;
 KW immunological disorder; arthritis; psoriasis; congenital heart defect;
 KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
 XX
 OS Homo sapiens.
 PN WO200257460-A2.
 XX
 PD 25-JUN-2002.
 XX
 PR 20-DEC-2001; 2001WO-US050459.
 XX
 PR 20-DEC-2000; 2000US-0256868P.
 PR 30-MAR-2001; 2001US-0280186P.
 PR 01-MAY-2001; 2001US-0287735P.
 PR 05-JUN-2001; 2001US-0295848P.
 PR 25-JUN-2001; 2001US-0300465P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
 PI Siemens N, Bol D, Schieven G, Finger J, Todderud CG, Basolino D;
 PI Krystek S, Mcatee P, Suchard S, Banaas D;
 DR WP1; 2002-599721/64.
 DR P-PSDB; ABR52407.
 XX
 PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 XX
 PS Example 7; Fig 19; 801pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid comprising a
 CC polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotides fully defined in the specification. The polynucleotide of
 CC the invention has antiproliferative, hepatotropic, nephrotropic,
 CC antiarthritic, antipneumatic, cardiant, and cytostatic activity. The
 CC polynucleotide may have a use in gene therapy. A polynucleotide or
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention
 XX
 SQ Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;

Query Match		100.0%;	Score 1995;	DB 6;	Length 5450;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1995;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGCCCATGATGATGATTTGGA	ACTCAATTTGTTACTGAGAGTTGATGCTCTGCTGAA	60	
DB	538	ATGCCCATGATGATGATTTGGA	ACTCAATTTGTTACTGAGAGTTGATGCTCTGCTGAA	597	
QY	61	AGTGGAAACGAAAAAGTCGCT	CTAATTGATAGCCGGCCATTTGTGGAAATCAATACATCC	120	
DB	598	AGTGGAAACGAAAAAGTCGCT	CTAATTGATAGCCGGCCATTTGTGGAAATCAATACATCC	657	
QY	121	CACATTTTGGAAAGCCATTA	TATCAACTGCTCCCAAGCTTATGAAAGCAAGGTTGCAACAG	180	
DB	658	CACATTTTGGAAAGCCATTA	TATCAACTGCTCCCAAGCTTATGAAAGCAAGGTTGCAACAG	717	
QY	181	GACAAAGTGTTAATTA	CAGAGCTCATCAGACATTCACGAAACATAAGTTGACATTGAT	240	
DB	718	GACAAAGTGTTAATTA	CAGAGCTCATCAGACATTCACGAAACATAAGTTGATGATGAT	777	
QY	241	TGCAGTCAGAAAGTTGAT	AGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA	300	
DB	778	TGCAGTCAGAAAGTTGAT	AGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA	837	
QY	301	GACGCTTTTCTCACTG	TACTCTGGGGTAACTGGAGAAAGGCTTCAACTCTGTTACCTG	360	
DB	838	GACGCTTTTCTCACTG	TACTCTGGGGTAACTGGAGAAAGGCTTCAACTCTGTTACCTG	897	
QY	361	CTTGCAGATGGGATTTGCT	GAGATTCTCGATGTGTTTCCCTGAGCCCTGTGAAAGAAATTC	420	
DB	898	CTTGCAGATGGGATTTGCT	GAGATTCTCGATGTGTTTCCCTGAGCCCTGTGAAAGAAATTC	957	
QY	421	ACTCTAGTCCCTAC	CTGSCATTTCTCAGCCTTGCTTACCTGTGTCACATTTGGGCCAAC	480	
DB	958	ACTCTAGTCCCTAC	CTGSCATTTCTCAGCCTTGCTTACCTGTGTCACATTTGGGCCAAC	1017	
QY	481	CGAATTTCTTCCCAATCT	TTTATCTTGTCGCTCCAGGGAATGTCTTCACAAAGAGCTGATA	540	
DB	1018	CGAATTTCTTCCCAATCT	TTTATCTTGTCGCTCCAGGGAATGTCTTCACAAAGAGCTGATA	1077	
QY	541	CAGCAGAAATGGGATTTG	TTATGTGTAAATGCGCAGCTATACCTGTCCAAAGCCTGACTTT	600	
DB	1078	CAGCAGAAATGGGATTTG	TTATGTGTAAATGCGCAGCTATACCTGTCCAAAGCCTGACTTT	1137	
QY	601	ATCCCGCAGTCTCA	TTTCTCTGCGCTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG	660	
DB	1138	ATCCCGCAGTCTCA	TTTCTCTGCGCTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG	1197	
QY	661	CCGTGTGTTGGACA	AAATCAGTAGATTTCAATTGAGAAAACAAAGCCTCCAAATGATGTGTT	720	
DB	1198	CCGTGTGTTGGACA	AAATCAGTAGATTTCAATTGAGAAAACAAAGCCTCCAAATGATGTGTT	1257	
QY	721	CTACTGCACTGTTTAG	CTGGGGAATCCCGGCTCCGSCACATCGCTATCGGCTACATCATG	780	
DB	1258	CTACTGCACTGTTTAG	CTGGGGAATCCCGGCTCCGSCACATCGCTATCGGCTACATCATG	1317	
QY	781	AAGAGATGACATG	TCCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACCTACT	840	
DB	1318	AAGAGATGACATG	TCCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACCTACT	1377	
QY	841	ATATCTCCAAACTT	CAATTTTCTGCGGCCAACTCTCTGACATATGAGAAAGATTAGAC	900	
DB	1378	ATATCTCCAAACTT	CAATTTTCTGCGGCCAACTCTCTGACATATGAGAAAGATTAGAC	1437	
QY	901	CAGCTGGAGATCAG	GGCCAAAGAGCAAACTCAAGCTGCTGCACTCTGGAGAAAGATTAGAC	960	
DB	1438	CAGCTGGAGATCAG	GGCCAAAGAGCAAACTCAAGCTGCTGCACTCTGGAGAAAGATTAGAC	1497	
QY	961	GAACTGTCTCCCTG	CTGTCTCAGAGGGTGGACAGAAAAACGAGACGCCCTCACTGCAACC	1020	
DB	1498	GAACTGTCTCCCTG	CTGTCTCAGAGGGTGGACAGAAAAACGAGACGCCCTCACTGCAACC	1557	
QY	1021	TGTGCGCACTCTG	CTACTCAGAGGCAAGACCAAAAGCCCGTGATCCCGCAGCGTG	1080	

Db	1558	TTGTCCGACTCTGCTACTCTCAGAGGACAGACGAAAGAGCCCGTGCATCCCGCAGCGTG	1617
Qy	1081	CCGACGCTGCCGAGCGTGCAGCCGCTCGCTGTATGAGAGACAGCCCGCTGTACAGCGCTC	1140
Db	1618	CCGACGCTGCCGAGCGTGCAGCCGCTCGCTGTATGAGAGACAGCCCGCTGTACAGCGCTC	1677
Qy	1141	AGTGGGCTGCACCTGTCCGAGACAGGCTGTGAAGACAGCAATTAAGCTTAAGCGTTC	1200
Db	1678	AGTGGGCTGCACCTGTCCGAGACAGGCTGTGAAGACAGCAATTAAGCTTAAGCGTTC	1737
Qy	1201	TCTCTGATTCAAATCAGTTTCAATTCAGCCAGATGGAGATCCCTTAATGAGCTC	1260
Db	1738	TCTCTGATTCAAATCAGTTTCAATTCAGCCAGATGGAGATCCCTTAATGAGCTC	1797
Qy	1261	TCCTCATCAAGAAAGTGGTTTGAAATATCAAACTTCCACTACTCTGATGGAGCCAC	1320
Db	1798	TCCTCATCAAGAAAGTGGTTTGAAATATCAAACTTCCACTACTCTGATGGAGCCAC	1857
Qy	1331	AAGCTATGCCAGTTCTTCCTGTTTCAGAACTATGAGACAGACTCCCGAAACGAGTCT	1380
Db	1858	AAGCTATGCCAGTTCTTCCTGTTTCAGAACTATGAGACAGACTCCCGAAACGAGTCT	1917
Qy	1381	GATTAAGAGAGAGACGACGATCCCGAAAGAGCTGAGACCCGACAGCCCTTACAGACGAG	1440
Db	1918	GATTAAGAGAGAGACGACGATCCCGAAAGAGCTGAGACCCGACAGCCCTTACAGACGAG	1977
Qy	1441	AGCAGCGATTCGATTCGGTGCAGAAACAGACAGAGTGGACCGGCGAGAGTCCCTTTA	1500
Db	1978	AGCAGCGATTCGATTCGGTGCAGAAACAGACAGAGTGGACCGGCGAGAGTCCCTTTA	2037
Qy	1501	TCCTCATGTCATGAAAGTGGAGCGGTGGAGAGCAATTAACACACAGCTTCCTTTTGGC	1560
Db	2038	TCCTCATGTCATGAAAGTGGAGCGGTGGAGAGCAATTAACACACAGCTTCCTTTTGGC	2097
Qy	1561	CTTTTCACACAGCCAGACGACCTTACGAAAGTCTGCTGGGCTTTAAGGGCTGGAC	1620
Db	2098	CTTTTCACACAGCCAGACGACCTTACGAAAGTCTGCTGGGCTTTAAGGGCTGGAC	2157
Qy	1621	TCGATATCTTGGCCCCCAGACCCCTTAACCCCTTCCCTGACACAGCTGTATTTGGC	1680
Db	2158	TCGATATCTTGGCCCCCAGACCCCTTAAACCCCTTCCCTGACACAGCTGTATTTGGC	2217
Qy	1681	AAGAGTCTCTACACATTCTTACTCTGCTCAAGCCATCTACGGAGGAGTGCAGATTCT	1740
Db	2218	AAGAGTCTCTACACATTCTTACTCTGCTCAAGCCATCTACGGAGGAGTGCAGATTCT	2277
Qy	1741	GCTTACAGCTCAGCCAGCTGCCCATCTGGCGGAGACAAAGTCTATTTCTGTGCGACGCG	1800
Db	2278	GCTTACAGCTCAGCCAGCTGCCCATCTGGCGGAGACAAAGTCTATTTCTGTGCGACGCG	2337
Qy	1801	CAGAAAGCCAAATGACAGAGCTGACTCGCGCGAGAGCTGCATGAAGAGAGCCCTTGAA	1860
Db	2338	CAGAAAGCCAAATGACAGAGCTGACTCGCGCGAGAGCTGCATGAAGAGAGCCCTTGAA	2397
Qy	1861	AAGCAGTTTAAACSCAGAGACTSCAAATTGGAAATTTGAGAGAGCATATGTCAAGAAC	1920
Db	2398	AAGCAGTTTAAACSCAGAGACTSCAAATTGGAAATTTGAGAGAGCATATGTCAAGAAC	2457
Qy	1921	AGGTCAAGGGAAGAGCTGTGGGAAAAGTGGGACATCAGTCTTATTTTGGGACAGATGAA	1980
Db	2458	AGGTCAAGGGAAGAGCTGTGGGAAAAGTGGGACATCAGTCTTATTTTGGGACAGATGAA	2517
Qy	1981	ATCATTTGAGGTCTCC	1995
Db	2518	ATCATTTGAGGTCTCC	2532

DJ		WT	18-JUN-2002 (first entry)
DE		XX	cDNA sequence of human (dual specificity phosphatase) DUSP-10.
KM		XX	Human; dual specificity phosphatases; DUSP-10; cancer; epilepsy; stroke;
KW		XX	neuronal degeneration syndrome; Alzheimer's disease; depression;
KX		XX	sclizophrenia; asthma; immune disorder; inflammatory process; arthritis;
OS		XX	osteoporosis; diabetes; gene therapy; chromosome 12; ss.
FH		XX	Homo sapiens.
FT		XX	Key Location/Qualifiers CDS 127..2124 /*tag= a /product= "DUSP-10 protein"
PN		XX	WO2001/77340-A1.
PD		XX	18-OCT-2001.
PF		XX	06-APR-2001; 2001MO-BPO03966.
PR		XX	10-APR-2000; 2000EP-00107143. (MERCK) MERCK PATENT GMBH. Duecker K; WIPI : 2002-010917/01. P-PSTDB ; AAU09946.
PT		XX	Newel dual specificity phosphatase polypeptides useful for treating cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's disease, depression, schizophrenia, asthma and immune disorders.
PS		XX	Claim 5; Page 34-37; 43pp; English.
CC		XX	The present invention relates to a new isolated dual specificity phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid sequence that is fully defined in the specification. The invention also provides a sequence encoded by a 3059 nucleotide sequence fully defined in the specification, and a sequence having at least 95 % identity to the polypeptide, or fragments or variants of DUSP-10. The invention is useful for treating cancer e.g. leukaemia, colon carcinoma, lung cancer, prostate cancer, metastasis of tumor cells, neo-angiogenesis, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's disease, depression, schizophrenia, cardiac myopathies, asthma, immune disorders, inflammatory processes e.g. arthritis, bowel disease, type I diabetes, osteoporosis, diabetes and diabetes associated diseases. The molecules of the invention are also useful as vaccines for inducing immunological agonistic or antagonistic compounds. Other uses of the invention include identifying membrane bound or soluble receptors, as a diagnostic reagent, in chromosome localisation studies, and as a valuable tool in tissue expression studies. The present sequence represents cDNA of the human dual specificity phosphatase, DUSP-10. DUSP-10 is located on chromosome 12
SQ		XX	Sequence 3059 BP; 831 A; 761 C; 709 G; 758 T; 0 U; 0 Other;
OY		YY	Query Match 99.8%; Score 1991.8; DB 6; Length 3059;
Db		YY	Best Local Similarity 99.9%; Pred. No. 0;
Mtches		YY	Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
I		ATGGGCCCATGAGATGATTGTGAACACTGCATAATTGTTA CTGAGA GTTGTCCTCGTGGAA	60
127		A TTGGCCAATGAGATGATTGTGAACCACTCAATTGTTA CTGAGA GGTTGGTCCTCGTGGAA	186
61		AGTGAACAAGAAAAGTCTCTCTAATTGATGAGCGGCGCATTTGTGGAATCAATCATGC	120
187		AGTGAACAAGAAAAGTCTCTCTCTAATTGATGAGCGGCGCATTTGTGGAATCAATCATGC	120

QY 121 CACATTTGGAGCCATTAAATCACTGCTCCAGGTTATGAGAGGAGTTGCAACAG 180
DB 247 CACATTTGGAGCCATTAAATCACTGCTCCAGGTTATGAGAGGAGTTGCAACAG 306
QY 181 GACAAAGTGTAAATTCAGAGCTCATCAGCAATTCAGGAAACATTAAGTTGACATTGAT 240
DB 307 GACAAAGTGTAAATTCAGAGCTCATCAGCAATTCAGGAAACATTAAGTTGACATTGAT 366
QY 241 TGCAGTACAGAGGTGTAGTTACAGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 300
DB 367 TGCAGTACAGAGGTGTAGTTACAGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 426
QY 301 GACTGTTTCTCACTGACTTCTGAGTAACTGGAAGAGCTTCAACTCTGTTCACTG 360
DB 427 GACTGTTTCTCACTGACTTCTGAGTAACTGGAAGAGCTTCAACTCTGTTCACTG 486
QY 361 CTTCGAGGTGGGTTTGTGAGTCTCTGTTGTTCCCTGAGCTCTGTGAAGAAATCC 420
DB 487 CTTCGAGGTGGGTTTGTGAGTCTCTGTTGTTCCCTGAGCTCTGTGAAGAAATCC 546
QY 421 ACTCTAGTCCCTTACCTGCAATTTCTCAGCTTGTCTTACCTGTGCAACATTTGGCCCAAC 480
DB 547 ACTCTAGTCCCTTACCTGCAATTTCTCAGCTTGTCTTACCTGTGCAACATTTGGCCCAAC 606
QY 481 GGAATTCCTCCCAATCTTTATCTTGTGCTGAGAGAGATGCTCCCAAGAGCTGATA 540
DB 607 GGAATTCCTCCCAATCTTTATCTTGTGCTGAGAGAGATGCTCCCAAGAGCTGATA 666
QY 541 CAGCAGAAATGGAGTTGTTATGTTTAAATGCGAGCTTATCTGTCCAAAGCCCTGACTTT 600
DB 667 CAGCAGAAATGGAGTTGTTATGTTTAAATGCGAGCTTATCTGTCCAAAGCCCTGACTTT 726
QY 601 ATCCCGGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 727 ATCCCGGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
QY 661 CCGTGTGTCAGCAATAGTATGATTTTCTGAGGAGCAAAAGGCTCCCAATGAGTGTG 720
DB 787 CCGTGTGTCAGCAATAGTATGATTTTCTGAGGAGCAAAAGGCTCCCAATGAGTGTG 846
QY 721 CTAGTGCATGTTTGTGCTGAGATCTCCGCTCCGCAACATGCTATGCTTACATCATG 780
DB 847 CTAGTGCATGTTTGTGCTGAGATCTCCGCTCCGCAACATGCTATGCTTACATCATG 906
QY 781 AAGAGATGCACTGCTTTTATGATGAAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
DB 907 AAGAGATGCACTGCTTTTATGATGAAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 966
QY 841 ATATCTCCAAACTTCAATTTTCTGAGGCACTCTGAGCTATGAGAGAAATTAAGAAC 900
DB 967 ATATCTCCAAACTTCAATTTTCTGAGGCACTCTGAGCTATGAGAGAAATTAAGAAC 1026
QY 901 CAGAGTGCAGATCAGGAGCCAAAGAGCAAACTCAAGCTGCTGCACTGAGAGAACCAAT 960
DB 1027 CAGAGTGCAGATCAGGAGCCAAAGAGCAAACTCAAGCTGCTGCACTGAGAGAACCAAT 1086
QY 961 GAACTGTCTCTGCTGTCTAAGAGGTGAGACAGAAAGCAAGAGCCCTCTCATGCCACC 1020
DB 1087 GAACTGTCTCTGCTGTCTAAGAGGTGAGACAGAAAGCAAGAGCCCTCTCATGCCACC 1146
QY 1021 TGTGCGCAGCTCTGTCTAAGAGGTGAGACAGAAAGGCGCTGATCCCGCAGAGGTG 1080
DB 1147 TGTGCGCAGCTCTGTCTAAGAGGTGAGACAGAAAGGCGCTGATCCCGCAGAGGTG 1206
QY 1081 CCGAGCTGCGCAGCTGAGCCGCTGCTGTAGAGAGACGCCGCTGTATCAGGCGCTC 1140
DB 1207 CCGAGCTGCGCAGCTGAGCCGCTGCTGTAGAGAGACGCCGCTGTATCAGGCGCTC 1266
QY 1141 AGTGGCTGCACTGCTGCTGAGAGAGCTGAGAGACAGCAATTAAGCTTCAAGCTTCTTC 1200
DB 1267 AGTGGCTGCACTGCTGCTGAGAGAGCTGAGAGACAGCAATTAAGCTTCAAGCTTCTTC 1326
QY 1201 TCTCTGATATCAAAATCAGTTTATATTCAGCCAGCATGAGAGATCTTATCAATGCTTC 1260

DB 1327 TCTCTGATATCAAAATCAGTTTATATTCAGCCAGATGAGACATCTTATCATGCTTC 1386
QY 1261 TCTCTATCAGAAAGTCTTTGAAATCTTCAAACTTCCACTCTCTGTGATGAGCAAC 1320
DB 1387 TCTCTATCAGAAAGTCTTTGAAATCTTCAAACTTCCACTCTCTGTGATGAGCAAC 1446
QY 1321 AAGCTATGCGAGTCTTCCCTGTTTCAAGAACTATCGGAGAGACTCCCGAAACAGTCT 1380
DB 1447 AAGCTATGCGAGTCTTCCCTGTTTCAAGAACTATCGGAGAGACTCCCGAAACAGTCT 1506
QY 1381 GATTAAGAGAGAACCCAGATCTCCCAAGAACTGAGACCGCCAGGCTTTCAGACAGCAG 1440
DB 1507 GATTAAGAGAGAACCCAGATCTCCCAAGAACTGAGACCGCCAGGCTTTCAGACAGCAG 1566
QY 1441 AGCAAGCATTTGATTTGCTGCTGAGAACCCAGAGAGTGCACCGCCCAAGAGTCCCTTTA 1500
DB 1567 AGCAAGCATTTGATTTGCTGCTGAGAACCCAGAGAGTGCACCGCCCAAGAGTCCCTTTA 1626
QY 1501 TCTCCACTGATGAGAGTGGGAGCGTGGAGGACAAATTAACAACAAGCTTCTTTGCGC 1560
DB 1627 TCTCCACTGATGAGAGTGGGAGCGTGGAGGACAAATTAACAACAAGCTTCTTTGCGC 1686
QY 1561 CTTTCCACAGCAGCAGAGCACTCAGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
DB 1687 CTTTCCACAGCAGCAGAGCACTCAGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1746
QY 1621 TCGGATATCTTGGCCCCCAGACCTTCACTCCCTTCCCTGACAGAGCTGATTTTGGC 1680
DB 1747 TCGGATATCTTGGCCCCCAGACCTTCACTCCCTTCCCTGACAGAGCTGATTTTGGC 1806
QY 1681 ACAGATCTCTCAACTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
DB 1807 ACAGATCTCTCAACTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1866
QY 1741 GCCTACAGCTGACAGCAGCTGCCACTGTCGAGAGCAAAAGTCTTCTGTGCGCAGGCG 1800
DB 1867 GCCTACAGCTGACAGCAGCTGCCACTGTCGAGAGCAAAAGTCTTCTGTGCGCAGGCG 1926
QY 1801 CAGAGCCAAAGTACAGAGCTGATCTGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB 1927 CAGAGCCAAAGTACAGAGCTGATCTGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1986
QY 1861 AAGCAGTTTAAAGCAGAGCTGCAATGGAATTTGAGAGAGCATCATGTAGAGAAC 1920
DB 1987 AAGCAGTTTAAAGCAGAGCTGCAATGGAATTTGAGAGAGCATCATGTAGAGAAC 2046
QY 1921 AGGTCAAGGAGAGCTGAGGAGAGTGGCAGTCACTTGTAGCTTTTCCGAGCAGATGAA 1980
DB 2047 AGGTCAAGGAGAGCTGAGGAGAGTGGCAGTCACTTGTAGCTTTTCCGAGCAGATGAA 2106
QY 1981 ATCATTGAGGTCTCC 1995
DB 2107 ATCATTGAGGTCTCC 2121

RESULT 4
ABK47596
ID ABK47596 standard; cDNA; 3496 BP.
XX
XX ABK47596;
AC
DT 02-JUL-2002 (first entry)
XX
XX
DE cDNA encoding human dual-specificity phosphatase-3 (DSP-16) protein.
KW Human: dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW cancer; graft-versus-host disease; allergy; metabolic disease;
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
KW intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; 88.
XX

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 562..2559
 FT /tag= a
 FT /product= "Human dual-specificity phosphatase-3 (DSP-16)
 protein"
 XX MO200226997-A2.
 XX
 XX 04-APR-2002.
 XX
 XX 25-SEP-2001; 2001WO-US030124.
 XX
 XX 26-SEP-2000; 2000US-0235487P.
 XX
 XX (CEPT)- CEPTYR INC.
 XX
 XX Luche RM, Wei B;
 XX
 XX WPI: 2002-315802/35.
 XX
 XX P-PSDB; AAU79156.
 XX
 XX New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 XX
 XX Claim 7; Fig 1; 87pp; English.
 XX
 XX The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC phosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity, for modulation of a proliferative response in a cell, survival
 CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present nucleic
 CC acid sequence is that of the human dual-specificity phosphatase-3 (DSP-
 CC 16) gene located on chromosome 12p. This sequence encodes the human DSP-
 CC 16 protein of the invention
 XX
 SQ Sequence 3496 BP; 939 A; 854 C; 856 G; 847 T; 0 U; 0 Other;
 Query Match 99.8%; Score 1991.8; DB 6; Length 3496;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 862 GACGTTTTCTACCTACTCTCTGAGTAACCTGAGAAAGACCTTCAACTCTGTTCACTG 921
 QY 361 CTTCGACAGTGGGTTTGTCTGAGTTCTCTGTTGTTTCCCTGGCCCTCTGTGAAGAAATCC 420
 DB 922 CTTCGACAGTGGGTTTGTCTGAGTTCTCTGTTGTTTCCCTGGCCCTCTGTGAAGAAATCC 981
 QY 421 ACTCTAGTCCCTACCTGCACTTCTCAGCTTGTCTTACCTGTGCCAACAATGGGCCAAC 480
 DB 982 ACTCTAGTCCCTACCTGCACTTCTCAGCTTGTCTTACCTGTGCCAACAATGGGCCAAC 1041
 QY 481 CGAATTTCTCCCAATCTTATCTTGTGCTGCGACGAGATGTCTCAACAAGAGCTGATG 540
 DB 1042 CGAATTTCTCCCAATCTTATCTTGTGCTGCGACGAGATGTCTCAACAAGAGCTGATG 1101
 QY 541 CAGCAGATGGGATTTGGTTAATGTTAAATGCCAGTAACTCTGTCCAAAGCTGACTTT 600
 DB 1102 CAGCAGATGGGATTTGGTTAATGTTAAATGCCAGTAACTCTGTCCAAAGCTGACTTT 1161
 QY 601 ATCCCGAGTCTCACTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 660
 DB 1162 ATCCCGAGTCTCACTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 1221
 QY 661 CCGTGTGTGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCAATGATGTGTT 720
 DB 1222 CCGTGTGTGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCAATGATGTGTT 1281
 QY 721 CTGATGACCTGTAGCTGGGATCTCCGCTCCGCAACATCGCTATGCGCTCATGATG 780
 DB 1282 CTGATGACCTGTAGCTGGGATCTCCGCTCCGCAACATCGCTATGCGCTCATGATG 1341
 QY 781 AAGAGGATGACATGCTTTAGATGAATGAGCTTACAGATTTGTGAAGAAAGAAACCTTACT 840
 DB 1342 AAGAGGATGACATGCTTTAGATGAATGAGCTTACAGATTTGTGAAGAAAGAAACCTTACT 1401
 QY 841 ATATCTCCAAATCTCAATTTCTGGCCCACTCTGTGACTATGAGAAAGATTAGAAC 900
 DB 1402 ATATCTCCAAATCTCAATTTCTGGCCCACTCTGTGACTATGAGAAAGATTAGAAC 1461
 QY 901 CAGACTGAGGATGAGGCGCAAGAGCAAACTCACTGTGACTGTGAGAGCCCAAT 960
 DB 1462 CAGACTGAGGATGAGGCGCAAGAGCAAACTCACTGTGACTGTGAGAGCCCAAT 1521
 QY 961 GAACCTGCTGCTGTCTCAAGGTTGACAGAAAGCGACAGCCCTCAGTCCAGCC 1020
 DB 1522 GAACCTGCTGCTGTCTCAAGGTTGACAGAAAGCGACAGCCCTCAGTCCAGCC 1080
 QY 1021 TGTCCGACTCTGTCTCACTCAAGGAGGAGGCAAAAGCCGTGATCCGCAAGCGTG 1080
 DB 1582 TGTCCGACTCTGTCTCACTCAAGGAGGAGGCAAAAGCCGTGATCCGCAAGCGTG 1641
 QY 1081 CCCAGCGTCCGAGCGCTGACAGCCGCTGCTTTAGAGACAGCCCGCTGTACAGGCGCTG 1140
 DB 1642 CCCAGCGTCCGAGCGCTGACAGCCGCTGCTTTAGAGACAGCCCGCTGTACAGGCGCTG 1701
 QY 1141 AGTGGCTGACCTGTCCGACAGAGGCTGTGAAGACAGCAATAGCTTCAAGCTTCTTC 1200
 DB 1702 AGTGGCTGACCTGTCCGACAGAGGCTGTGAAGACAGCAATAGCTTCAAGCTTCTTC 1260
 QY 1201 TCTCTGATATCAATCTGTTTCAATTTCAAGCCAGATGAGCAATCTTACATGCTTTC 1260
 DB 1762 TCTCTGATATCAATCTGTTTCAATTTCAAGCCAGATGAGCAATCTTACATGCTTTC 1821
 QY 1261 TCTCTGATATCAATCTGTTTCAATTTCAAGCCAGATGAGCAATCTTACATGCTTTC 1320
 DB 1822 TCTCTGATATCAATCTGTTTCAATTTCAAGCCAGATGAGCAATCTTACATGCTTTC 1381
 QY 1321 AAGCTATGCGCTTCTCCCTGTTGAGAACTATGTGAGACAGATCTCCGAAACAGTCTCT 1380
 DB 1882 AAGCTATGCGCTTCTCCCTGTTGAGAACTATGTGAGACAGATCTCCGAAACAGTCTCT 1941
 QY 1381 GATTAAGAGGAGGCAATCCCAAGAGCTGACAGCCGAGGCTTTCAGACGCCAG 1440

Db	1942	GATTAAGGAGAAGCAGCATCCCAAGAAAGCTGGACGCCAGAGCCCTTCAGACAGCAG	2001
Qy	1441	AGCAAAGCATTTGCATTCCGTCAGAACCAGCAGCAGTGGCACCGCCAGAGGTCCCTTTTA	1500
Db	2002	AGCAAGGCATTGGACTTCGGTCAGAACCAAGCAGCAGTGGCACCGGCCAAGAGTCCCTTTTA	2061
Qy	1501	TCTCCACTGCATCGAAGTGGAGGGGTGGAGACAATTACGACACAGCTTCCTTTTGCGC	1566
Db	2062	TCTCCACTGCATCGAAGTGGAGGGGTGGAGACAATTACGACACAGCTTCCTTTTGCGC	2121
Qy	1561	CTTTTCACACGACCCAGCAGCACCTCAACGAAATCTGTGGCCCTGGGCTTTAAAGGGCTGGCAC	1620
Db	2122	CTTTTCACACGACCCAGCAGCACCTCAACGAAATCTGTGGCCCTGGGCTTTAAAGGGCTGGCAC	2181
Qy	1621	TCGGATATCTTTGGCCCCCCCAGAACCTTACCCCTTCCCTGACCAAGCAGTGGTATTTTGCC	1680
Db	2182	TCGGATATCTTTGGCCCCCCCAGAACCTTACCCCTTCCCTGACCAAGCAGTGGTATTTTGCC	2241
Qy	1681	ACAGAAGTCTACACTTCTACTCTGCGCTCAACCATCTACGGAAGGAGGAGCCAGTTACTCT	1740
Db	2242	ACAGAAGTCTACACTTCTACTCTGCGCTCAACCATCTACGGAAGGAGGAGCCAGTTACTCT	2301
Qy	1741	GCCTACAGCTGCACCCAGCTCCCACTTGGCGGAGACCAAGTCTATTCTGTGCGCAGGCGG	1800
Db	2302	GCCTACAGCTGCACCCAGCTCCCACTTGGCGGAGACCAAGTCTATTCTGTGCGCAGGCGG	2361
Qy	1801	CAGAAGCCCAATGACAGAGCTGACTCGCGCGGAGAGCTGGCATGAAGAGAGCCCTTTGAA	1860
Db	2362	CAGAAGCCCAATGACAGAGCTGACTCGCGCGGAGAGCTGGCATGAAGAGAGCCCTTTGAA	2421
Qy	1861	AAGCAGTTTAAACCCAGAAAGCTGGCCMAATGGAAITTTGGANAGAGCATATGTCAAGAAC	1920
Db	2422	AAGCAGTTTAAACCCAGAAAGCTGGCCMAATGGAAITTTGGANAGAGCATATGTCAAGAAC	2481
Qy	1921	AGGTCAACGGGAAGAGCTGGGGGAAAGTGGGGCAGTCACTTAGCTTTTGGCGGAGCATGGAA	1980
Db	2482	AGGTCAACGGGAAGAGCTGGGGGAAAGTGGGGCAGTCACTTAGCTTTTGGCGGAGCATGGAA	2541
Qy	1981	ATCATTTGAGGTCTCC 1995	
Db	2542	ATCATTTGAGGTCTCC 2556	
<hr/>			
RESULT 5			
XX	AASI4639	standard; cDNA; 3544 BP.	
XX	AASI4639;		
XX	AC		
XX	AD		
XX	DT	18-DEC-2001 (first entry)	
XX	XX		
DE	Human CDNA encoding dual specificity phosphatase 21117.		
KW	Human; sp; dual specificity phosphatase 21117; hepatotropic; cytosolic;		
KW	hematopoietic disorder; autoimmune disorder; diabetes mellitus;		
KW	rheumatoid arthritis; multiple sclerosis; Crohn's disease;		
KW	liver disorder; erythroid associated disorder; haemolytic anaemia;		
KW	cellular proliferative; differentiative disorder; leukaemia;		
KW	acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer.		
XX	OS	Homo sapiens.	
XX	OS		
PH	Key	Location/Qualifiers	
PH	FT	589..2586	
PT	CDS	/**tag= a	
FT		/product= "Dual specificity phosphatase 21117"	
XX	MO200173059-A2.		
PN			
PD	04-OCT-2001.		
FP	23-MAR-2001; 2001WO-US009477.		
XX			

PR	24-MAR-2000; 2000US-0191858P.
XX	(MILL-) MILLENNIUM PHARM INC.
PA	Meyers RA;
XX	WPI; 2001-611635/70.
XX	P-PSDB; AAU09016.
DR	
XX	
FT	New human dual specificity polypeptides and nucleic acids for diagnosis
PT	of disease and treatment of e.g. liver disorders.
XX	
PS	Claim 1; Fig 1; 143pb; English.
XX	
CC	The invention relates to two novel human dual specificity phosphatases
CC	designated 21117 and 38692, the nucleic acids encoding them (including
CC	fragments, allelic variants, their complements or nucleic acids that
CC	hybridise to them) and antibodies raised against the proteins. The
CC	antibody is useful for detecting the presence of the polypeptide, and the
CC	nucleic acid fragments are useful for detecting the presence of the
CC	nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
CC	antisense sequences) are useful for modulating the activity or expression
CC	of the polypeptide/nucleic acid, useful for the treatment of e.g.
CC	hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
CC	rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
CC	listed in the specification) liver disorders, erythroid associated
CC	disorders (e.g. haemolytic anaemia) cellular proliferative or
CC	differentiative disorders, leukaemias (e.g. acute myeloid leukaemia),
CC	carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and
CC	38692 are also useful for modulating the proliferation, survival,
CC	migration or differentiation of a 38692 or 21117-expressing cell. The
CC	polypeptide and nucleic acids are useful for identifying modulating
CC	agents. The present sequence encodes the dual specificity phosphatase
CC	21117
XX	
XX	Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;
SO	
	Query Match 99.8%; Score 1991.8; DB 5; Length 3544;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY	1 ATGGCCCATGATGATGATGGAATCTCAATTGTACTAGAGCTTGTGGCTCTGCTGGA 60
DB	589 ATGGCCCATGATGATGATGGAATCTCAATTGTACTAGAGCTTGTGGCTCTGCTGGA 648
OY	61 AGTGAACGGAAAGAGTCGTCTAATTGATAGCGCGCATTTGGAAATCAATATACCC 120
DB	649 AGTGAACGGAAAGAGTCGTCTAATTGATAGCGCGCATTTGGAAATCAATATACCC 708
OY	121 CACATTTTGGAAAGCATTAATATCAACTGCTCCAAAGCTTATGAAACGAAAGTTGCCA 180
DB	709 CACATTTTGGAAAGCATTAATATCAACTGCTCCAAAGCTTATGAAACGAAAGTTGCCA 768
OY	181 GACAAAGTGTAAATTAACAGAGCTCATCAGATTCAGCGAAACATTAAGTTGACATTGAT 240
DB	769 GACAAAGTGTAAATTAACAGAGCTCATCAGATTCAGCGAAACATTAAGTTGACATTGAT 828
OY	241 TGCAGTCAAGAAAGTTGATAGTTTACAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
DB	829 TGCAGTCAAGAAAGTTGATAGTTTACAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 888
OY	301 GACGTGTTTCTCACTGTACTCTTGGGTAAATCTGAGAAAGAGCTTCAACTCTGTTCACCTG 360
DB	889 GACGTGTTTCTCACTGTACTCTTGGGTAAATCTGAGAAAGAGCTTCAACTCTGTTCACCTG 948
OY	361 CTTGCAAGTGGGTTTGTCTGAGTTCTCTCGTTGTTTCCCTGGCCTCTGTGAAGAAATCC 420
DB	949 CTTGCAAGTGGGTTTGTCTGAGTTCTCTCGTTGTTTCCCTGGCCTCTGTGAAGAAATCC 1008
OY	421 ACTGTAGTCCCTACCTGATTTTTCACGCTTGTACTCTGTGTGCAACATTTGGGCCAAC 480
DB	1009 ACTGTAGTCCCTACCTGATTTTTCACGCTTGTACTCTGTGTGCAACATTTGGGCCAAC 1068

QY 481 CGAATCTCCCAATCTTATCTTGCTGCCAGCAGATGCTCAACAAGAGCTGATA 540
 Db 1069 CGAATCTCTCCCAATCTTATCTTGCTGCCAGCAGATGCTCAACAAGAGCTGATA 1128
 QY 541 CAGCAAAATGGGATTTGTTATGTTAAATGCGAGTATACCTGTCCAAAGCTGACCTT 600
 Db 1129 CAGCAAAATGGGATTTGTTATGTTAAATGCGAGTATACCTGTCCAAAGCTGACCTT 1188
 QY 601 ATCCCGAGTCTCAATTTCCGCGCTGTGCTGTGAATGACGCTTTTGAGAGAAAATTTTG 660
 Db 1189 ATCCCGAGTCTCAATTTCCGCGCTGTGCTGTGAATGACGCTTTTGAGAGAAAATTTTG 1248
 QY 661 CCGTGTGGGACAAATAGTATGATTTGATGAAGAAAGCAAGCCCTCCAAATGATGTT 720
 Db 1249 CCGTGTGGGACAAATAGTATGATTTGATGAAGAAAGCAAGCCCTCCAAATGATGTT 1308
 QY 721 CTATGACATGTTTAAAGTGGATCTCCGCTCCGCAACCATGCTATGCTTACATCAAG 780
 Db 1309 CTATGACATGTTTAAAGTGGATCTCCGCTCCGCAACCATGCTATGCTTACATCAAG 1368
 QY 781 AAGAGATGGAATGCTTATGATGAGCTTACAGATTTGTGAAAGAAAAGACTTACT 840
 Db 1369 AAGAGATGGAATGCTTATGATGAGCTTACAGATTTGTGAAAGAAAAGACTTACT 1428
 QY 841 ATATCTCAAACTTCAATTTTCTGGGCCAATCTCTGATCTATGAAAGAAATTAAGAC 900
 Db 1429 ATATCTCAAACTTCAATTTTCTGGGCCAATCTCTGATCTATGAAAGAAATTAAGAC 1488
 QY 901 CAGATGAGAGATCAGGGCCAAAGAGAACTCAAGCGCTGACCTGTGAGAAAGCCAAAT 960
 Db 1489 CAGATGAGAGATCAGGGCCAAAGAGAACTCAAGCGCTGACCTGTGAGAAAGCCAAAT 1548
 QY 961 GAACCTGTCCCTGCTGTCTCAGAGGTGAGACAAAAGCAGACGCCCCCTCACTCCACCC 1020
 Db 1549 GAACCTGTCCCTGCTGTCTCAGAGGTGAGACAAAAGCAGACGCCCCCTCACTCCACCC 1608
 QY 1021 TGTGCGCACTGTGCTACTCAGAGCAGACAGACAAAGGCGGTGATCCCGCAGCGTG 1080
 Db 1609 TGTGCGCACTGTGCTACTCAGAGCAGACAGACAAAGGCGGTGATCCCGCAGCGTG 1668
 QY 1081 CCCAGCGTCCAGCAGTGCAGCGTGTGAGAGACAGCCCGCTGTGACAGCGCTC 1140
 Db 1669 CCCAGCGTCCAGCAGTGCAGCGTGTGAGAGACAGCCCGCTGTGACAGCGCTC 1728
 QY 1141 AGTGGGCTGCACTGTCTCCGAGACAGCGTGTGAAAGCAGCAATTAAGCTTCCCTTC 1200
 Db 1729 AGTGGGCTGCACTGTCTCCGAGACAGCGTGTGAAAGCAGCAATTAAGCTTCCCTTC 1788
 QY 1201 TCTCTGATATCAATCAATCATTTGATATGAGCCAGATGAGCAATCTTACATGCTTC 1260
 Db 1789 TCTCTGATATCAATCAATCATTTGATATGAGCCAGATGAGCAATCTTACATGCTTC 1848
 QY 1261 TCTCTATCAGAAAGATCTTTGGAATATCAAACTTCACTGTGAGTGGACCAAC 1320
 Db 1849 TCTCTATCAGAAAGATCTTTGGAATATCAAACTTCACTGTGAGTGGACCAAC 1908
 QY 1321 AAGCTATGCAATTTCTCCCTGTGAGAGAACTATGAGCAATCTCCGAAACCAATGCT 1380
 Db 1909 AAGCTATGCAATTTCTCCCTGTGAGAGAACTATGAGCAATCTCCGAAACCAATGCT 1968
 QY 1381 GATTAAGAGAAAGCAGATCCCAAGAAAGCTGACAGCCGCTTCAAGACAGCCAG 1440
 Db 1969 GATTAAGAGAAAGCAGATCCCAAGAAAGCTGACAGCCGCTTCAAGACAGCCAG 2028
 QY 1441 AGCAAGCGATTTGATTTGGTTCAGAAACAGACAGATGAGCAGCCGCAAGAGTCCCTTTTA 1500
 Db 2029 AGCAAGCGATTTGATTTGGTTCAGAAACAGACAGATGAGCAGCCGCAAGAGTCCCTTTTA 2088
 QY 1501 TCTCCATCTGATTCGAATGAGAGGTGAGAGCAATTAACAACAAGCTTCCCTTTTGGGC 1560
 Db 2089 TCTCCATCTGATTCGAATGAGAGGTGAGAGCAATTAACAACAAGCTTCCCTTTTGGGC 2148
 QY 1561 CTTTTCACGACGACGACGACCTCAAGAAAGTGTGCTGGCTTGAAGGGCTGGCAC 1620

Db 2149 CTTTTCACGACGACGACGACGACCTCAAGAAAGTGTGCTGGCTTGAAGGGCTGGCAC 2208
 QY 1621 TCGGATATCTTTGGCCCCCAGACTTCAACCCCTTCCCTGACACAGACCTGTAATTTTGGC 1680
 Db 2209 TCGGATATCTTTGGCCCCCAGACTTCAACCCCTTCCCTGACAGACCTGTAATTTTGGC 2268
 QY 1681 ACAGAGTCTCACTTCTACTGTGCTGAGCCATCTTAAGAGAGCTGCAATTTTGGC 1740
 Db 2269 ACAGAGTCTCACTTCTACTGTGCTGAGCCATCTTAAGAGAGCTGCAATTTTGGC 2328
 QY 1741 GCTTACAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
 Db 2329 GCTTACAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388
 QY 1801 CAGAAAGCAAGTGAACAGAGCTGACCTGCGCGGAGAGTGGACAGAAAGAGCCCTTTTGA 1860
 Db 2389 CAGAAAGCAAGTGAACAGAGCTGACCTGCGCGGAGAGTGGACAGAAAGAGCCCTTTTGA 2448
 QY 1861 AAGAGTTTAAACGAGAACTGCTCAATGAAATTTGAGAGAGCATATGTCAGAGAAC 1920
 Db 2449 AAGAGTTTAAACGAGAACTGCTCAATGAAATTTGAGAGAGCATATGTCAGAGAAC 2508
 QY 1921 AGGTCAAGGAGAGAGTGGGAAAGTGGGAGTCACTGCTTGTGGGACAGATGAA 1980
 Db 2509 AGGTCAAGGAGAGAGTGGGAAAGTGGGAGTCACTGCTTGTGGGACAGATGAA 2568
 QY 1981 ATCATTGAGTCTCC 1995
 Db 2569 ATCATTGAGTCTCC 2583

RESULT 6
 ABK49402
 ID ABK49402 standard; cDNA; 3544 BP.
 XX
 AC ABK49402;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE cDNA encoding human dual specificity phosphatase 21117 protein.
 XX
 KW Human; dual specificity phosphatase 21117; erythroid-related disorder;
 KW haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
 KW erythrocytosis; liver-related disorder; cancer; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 589..2586
 FT /*tag= a
 FT /product= "Human dual specificity phosphatase 21117"
 FT /note= "Specifically claimed in claim 2"
 PN US2002034807-A1.
 PD 21-MAR-2002.
 PP 23-MAR-2001; 2001US-00816494.
 PR 24-MAR-2000; 2000US-0191858P.
 PA (MEYE/) MEYERS R A.
 XX
 PI Meyers RA;
 DR WPI; 2002-351088/38.
 DR P-PSDB; AAU79929.
 XX
 PT New nucleic acids, designated 38692 and 21117, encoding dual specificity
 PT phosphatases for treating cell proliferation and differentiation
 PT disorders including hematopoietic and erythroid-related disorders and
 PT cancers.

Db	2449	AACGACGTTTAAACGCAAAAGCTGCCAAATGGAAATTTGGAGAGAGCATCATGTTCAGAGAAC	2508
QY	1921	AGGTCACGGGAAGACTGTGGGAAAGTGGCAGTCAGTCAGCTTTTCGGGACGATGGAA	1980
Db	2509	AGSTACCGGAAAGAACTGTGGGAAAGTGGGAGTCAGTCAGCTTTTCGGGACGATGGAA	2568
QY	1981	ATCATTTGAGGCTTCC	1995
Db	2569	ATCATTTGAGGCTTCC	2583

RESULT 7	
ABK14474	
ID	ABK14474 standard; cDNA; 3766 BP.
XX	

..._protease_/_ (pp7)

CC Huntington's disease, dementia or Parkinson's disease, developmental
 CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
 CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
 CC or sarcoma. The present nucleic acid sequence encodes human protein
 CC phosphatase 7 (P77) which is one of several human protein phosphatases
 CC (AA075783-AA075792) of the invention
 SX
 SQ Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;

Qy	301	CAGACTGAGGATCAAGGACCAGAAAGAGCAAACTCAAGCTGTGCACTTGAGAAAGCAAT	960
Db	1438	CAGACTGAGGATCAAGGACCAGAAAGAGCAAACTCAAGCTGTGCACTTGAGAAAGCAAT	1497
Qy	961	GAACCTGTCCCTGTCTCTCAAGAGGTGAGACAAAAGCGAGACGCCCTCACTCAACC	10320
Db	1498	GAACCTGTCCCTGTCTCTCAAGAGGTGAGACAAAAGCGAGACGCCCTCACTCAACC	1557
Qy	1021	TGTGCGCACTCTGCTACTCAAGGACAGAGACAAAGGCCGTGTCAATCCGACAGGTG	10860
Db	1558	TGTGCGCACTCTGCTACTCAAGGACAGAGACAAAGGCCGTGTCAATCCGACAGGTG	16171
Qy	1081	CCGACGCGCCAGGCTGACGCGGTGGTTAGAGACAGCGCGTGTGTCAAGCGCTG	11464
Db	1618	CCGACGCGCCAGGCTGACGCGGTGGTTAGAGACAGCGCGTGTGTCAAGCGCTG	16777
Qy	1141	AGTGGGCTGCACCTGTCCGAGACAGGCTGAGAGACAGCAATTAAGCTCAAGCGTCTTC	12004
Db	1678	AGTGGGCTGCACCTGTCCGAGACAGGCTGAGAGACAGCAATTAAGCTCAAGCGTCTTC	17373
Qy	1201	TCTCTGATATCAAAATCAGTTTCATATTCAGCCAGCATGGCAGCATCTTACATGGCTTC	12660
Db	1738	TCTCTGATATCAAAATCAGTTTCATATTCAGCCAGCATGGCAGCATCTTACATGGCTTC	17971
Qy	1261	TCCCTCATAGAAAGATGCTTGGAAATCTAACAAACCTTCACTACTCTGTGAGAGCAAC	13220
Db	1798	TCCCTCATAGAAAGATGCTTGGAAATCTAACAAACCTTCACTACTCTGTGAGAGCAAC	18572
Qy	1321	AAGCTATGCAAGTTCTCCCTGTTCAGAGAACTATGAGAGACAGCTCCGGAACAGTCT	13880
Db	1858	AAGCTATGCAAGTTCTCCCTGTTCAGAGAACTATGAGAGACAGCTCCGGAACAGTCT	19171
Qy	1381	GATTAAGAGAGAACCCAGCATCCCAAGAGCTGCAGACCGCCAGGCTTCAGACAGCCAG	14440
Db	1918	GATTAAGAGAGAACCCAGCATCCCAAGAGCTGCAGACCGCCAGGCTTCAGACAGCCAG	19777
Qy	1441	AGCAAGCGATTTGCAATTCGGTCAAGAACCAAGCAGCATGGCACCGCCAGAGTCCCTTTTA	15000
Db	1978	AGCAAGCGATTTGCAATTCGGTCAAGAACCAAGCAGCATGGCACCGCCAGAGTCCCTTTTA	20372
Qy	1501	TCTCCACTGCATCGAAGTGGAGAGGTGAGAGCAATTACCAACCAAGCTTCCTTTTGGGC	15660
Db	2038	TCTCCACTGCATCGAAGTGGAGAGGTGAGAGCAATTACCAACCAAGCTTCCTTTTGGGC	20972
Qy	1561	CTTTTCACACGACCAAGCACTCAAGAACTGTGTGACCTGTGGCTTTAAGGCTGGCAC	16220
Db	2098	CTTTTCACACGACCAAGCACTCAAGAACTGTGTGACCTGTGGCTTTAAGGCTGGCAC	21572
Qy	1621	TCCGATATCTTTGGCCCCCAGACCTCTACCCCTTCCTGACACAGCATGTGTATTTTGGC	16880
Db	2158	TCCGATATCTTTGGCCCCCAGACCTCTACCCCTTCCTGACACAGCATGTGTATTTTGGC	22171
Qy	1681	ACAGAGTCTCAACCTTACTCTGTGCTCAAGCATCTACGAGAGGAGCCAGTTACTCT	17440
Db	2218	ACAGAGTCTCAACCTTACTCTGTGCTCAAGCATCTACGAGAGGAGCCAGTTACTCT	22777
Qy	1741	GCTTACAGCTGACGACGACCTGCCACTTTCGCGAGACCAAGTCTTATCTGTGCGAGCGG	18000
Db	2278	GCTTACAGCTGACGACGACCTGCCACTTTCGCGAGACCAAGTCTTATCTGTGCGAGCGG	23372
Qy	1801	CAGAAAGCCAGTACAGAGCTGACTCGCGGCGAGAGCTGGCATGAGAGAGCCCTTTGAA	18660
Db	2338	CAGAAAGCCAGTACAGAGCTGACTCGCGGCGAGAGCTGGCATGAGAGAGCCCTTTGAA	23972
Qy	1861	AAGCAGTTTAAACGAGAGCTGCCAAATGGAATTTGAGAGAGCATCATTCAGAGAAC	19220
Db	2398	AAGCAGTTTAAACGAGAGCTGCCAAATGGAATTTGAGAGAGCATCATTCAGAGAAC	24572
Qy	1921	AGGTCAACGGAGAGAGCTGGGAGAAAGTGGGACAGTCACTTTTCGGGACAGATGAA	19880
Db	2458	AGGTCAACGGAGAGAGCTGGGAGAAAGTGGGACAGTCACTTTTCGGGACAGATGAA	25172
Qy	1981	ATCATTTAGGTCTCC	1995

Db	2518	ATCATTGAGTCTCC	2532
RESULT 8			
ABN83966			
ID	ABN83966	standard; DNA; 4790 BP.	
XX	ABN83966;		
AC			
XX	06-SEP-2002 (first entry)		
DT			
XX	Human gene sequence #13.		
DE			
XX	Human; brain; consil; hippocampus; foetal brain; diagnosis; gene; ds.		
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	184..2181	
XX		/*tag= a	
XX	MO200252005-A1.		
PN			
XX	04-JUL-2002.		
PD			
XX	20-DEC-2001; 2001WO-JP011217.		
PF			
XX	22-DEC-2000; 2000JP-00389742.		
PR			
XX	(KAZU-) KAZUSA DNA RES INST FOUND.		
PA	(CELE-) CELESTAR LEXICO-SCI LTD.		
XX			
XX	Ohara O, Nagase T, Nakajima D;		
PI			
XX	WPI: 2002-500762/53.		
DR	P-PSDB; ABB97946.		
XX			
PT	Genes and their expression products cloned from human cDNA libraries for		
XX	treatment and diagnosis of diseases associated with their expression.		
XX			
PS	Claim 1(a); Page 111-117; 238pp; Japanese.		
XX			
XX	The invention relates to DNA encoding polypeptides directly cloned from		
CC	cDNA libraries originating in adult whole brain, human consil, human		
CC	adult hippocampus and human foetal whole brain. Polypeptides and		
CC	polymucleotides of the invention may be used in the investigation of		
CC	differential expression of the DNA sequences in normal subjects and		
CC	disease patients. They may also be used in the production of antibodies,		
CC	oligonucleotide probes and DNA chips for diagnosis and identification of		
CC	drugs for treatment of diseases with which the DNA sequences are		
CC	associated. The sequences given in records ABN83954-ABN83984 represent		
CC	human gene sequences of the invention		
XX			
SQ	Sequence 4790 BP; 1293 A; 1064 C; 1152 G; 1281 T; 0 U; 0 Other;		
XX			
Query Match	99.8%; Score 1991.8; DB 6; Length 4790;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 1993; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
QY	1 ATGGCCCATGATGATGATTGGAATCTCAATTGTAAGAGAGTTGGTCTCTGCTGGAA	60	
DB	184 ATGGCCCATGATGATGATTGGAATCTCAATTGTAAGAGAGTTGGTCTCTGCTGGAA	243	
QY	61 AGTGGAAACGGAAGAAAGCTGCTTAATGATGATGACCGGCAATTTGGGAATCAATACATCC	120	
DB	244 AGTGGAAACGGAAGAAAGCTGCTTAATGATGATGACCGGCAATTTGGGAATCAATACATCC	303	
QY	121 CACATTTTGGAGGCAATTAATATCAATCTGCTCCAAAGCTTATGAAGCGAAGTTGCCAAG	180	
DB	304 CACATTTTGGAGGCAATTAATATCAATCTGCTCCAAAGCTTATGAAGCGAAGTTGCCAAG	363	
QY	181 GACAAAGCTTAAATTAACAGAGCTCATCAGCAATTCAGCGAAACATTAAGTTGACATTCAT	240	

Db 364 GACAAAGTGTATTAACAGAGCTCATTCCAGCAATTCAGGAAACATAAGTTGACATTGAT 423
 Qy 241 TGAAGTCAGAAAGGTTGATGTTTACGATCAAAAGCTCCAAAGTGTGCTCTCTCTCA 300
 Db 424 TGAAGTCAGAAAGGTTGATGTTTACGATCAAAAGCTCCAAAGTGTGCTCTCTCTCA 483
 Qy 301 GACTGTTTCTCACTGTAATCTGAGTAACTGAGAAAGCTTCAACTGTGTACCTG 360
 Db 484 GACTGTTTCTCACTGTAATCTGAGTAACTGAGAAAGCTTCAACTGTGTACCTG 543
 Qy 361 CTGAGAGTGGGTTTGTGAGTGTCTGTGTTTCCCTGGGCTCTGTAGAGAAATCC 420
 Db 544 CTGAGAGTGGGTTTGTGAGTGTCTGTGTTTCCCTGGGCTCTGTAGAGAAATCC 603
 Qy 421 ACTTAGTCCCTACCTGCAATTTCTCAGCCTTGTACCTGTTGCCAAATTTGGGCCAAC 480
 Db 604 ACTTAGTCCCTACCTGCAATTTCTCAGCCTTGTACCTGTTGCCAAATTTGGGCCAAC 663
 Qy 481 CGAATTTCTCCCAATCTTTATCTGTGCTGCGCAGCGAGATGTCTCAACAGAGAGTGATA 540
 Db 664 CGAATTTCTCCCAATCTTTATCTGTGCTGCGCAGCGAGATGTCTCAACAGAGAGTGATA 723
 Qy 541 CAGCAGAAATGGAAATGGTTATGTGTAAATGACAGTATACCTGTCCAAAGCTGACTTT 600
 Db 724 CAGCAGAAATGGAAATGGTTATGTGTAAATGACAGTATACCTGTCCAAAGCTGACTTT 783
 Qy 601 ATCCCGAGTCTATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 660
 Db 784 ATCCCGAGTCTATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 843
 Qy 661 CCGTGGTTGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCAATGATGTGT 720
 Db 844 CCGTGGTTGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCAATGATGTGT 903
 Qy 721 CTAGTGCATGTTTACGAGGATCTCCGCTCGGCCACATCCCTATCCCTACATCATG 780
 Db 904 CTAGTGCATGTTTACGAGGATCTCCGCTCGGCCACATCCCTATCCCTACATCATG 963
 Qy 781 AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAAAAGACCTTACT 840
 Db 964 AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAAAAGACCTTACT 1023
 Qy 841 ATATCTCCAACTTCATTTTCTGGGCAATCTCTGAGTATGAGAAAGATTAAAGAC 900
 Db 1024 ATATCTCCAACTTCATTTTCTGGGCAATCTCTGAGTATGAGAAAGATTAAAGAC 1083
 Qy 901 CAGACTGAGCATCAGGGCCAAAGACAATCAAGCTGTGSCACTCTGAGAAAGCCAAAT 960
 Db 1084 CAGACTGAGCATCAGGGCCAAAGACAATCAAGCTGTGSCACTCTGAGAAAGCCAAAT 1143
 Qy 961 GAACTGTCCCTGCTGTCTCAGAGGTGACAGAAAACGAGAGCCCTCAGTCCACC 1020
 Db 1144 GAACTGTCCCTGCTGTCTCAGAGGTGACAGAAAACGAGAGCCCTCAGTCCACC 1203
 Qy 1021 TGTGCGCATCTGTCTACCTCAGAGGACAGAGCAAAAGGCCCGTGCATCCCGCAGCGTG 1080
 Db 1204 TGTGCGCATCTGTCTACCTCAGAGGACAGAGCAAAAGGCCCGTGCATCCCGCAGCGTG 1263
 Qy 1081 CCCAGCGTCCAGAGGTGACAGCCGCTGTTAGAGACAGCCCGCTGTGACAGGGCTC 1140
 Db 1264 CCCAGCGTCCAGAGGTGACAGCCGCTGTTAGAGACAGCCCGCTGTGACAGGGCTC 1323
 Qy 1141 AGTGGGCTGACCTGTCCGACAGAGGCTGAAAGACAGCAATAGTCAAGGTTCTTTC 1200
 Db 1324 AGTGGGCTGACCTGTCCGACAGAGGCTGAAAGACAGCAATAGTCAAGGTTCTTTC 1383
 Qy 1201 TCTCTGATATCAATCATGTTTCAATATGACGAGATGGCAGATCTTATCATGGCTTC 1260
 Db 1384 TCTCTGATATCAATCATGTTTCAATATGACGAGATGGCAGATCTTATCATGGCTTC 1443
 Qy 1261 TCCCATCAAGAAATGCTTTGGAATCTAACACCTTCACTACTGAGAGGAGCAAC 1320

Db 1444 TCCATCATGAGAAATGCTTTGGAATCTAACAAACCTTCACTACTGATGGAGCAAC 1503
 Qy 1321 AAGCTATGCAAGTTCTCCCTGTGTAGAGACATATCGAGACAGATCTCCGAAACAGATCT 1380
 Db 1504 AAGCTATGCAAGTTCTCCCTGTGTAGAGACATATCGAGAGAGATCTCCGAAACAGATCT 1563
 Qy 1381 GATTAAGAGAAAGCCAGATCCCAAGAAAGCTGAGACCCGACGCTTTCAGACAGCCAG 1440
 Db 1564 GATTAAGAGAAAGCCAGATCCCAAGAAAGCTGAGACCCGACGCTTTCAGACAGCCAG 1623
 Qy 1441 AGCAAGCATTTGCAATGCTGTGAGAAACAGAGAGTGGCACCGCCAGAGATCCCTTTTA 1500
 Db 1624 AGCAAGCATTTGCAATGCTGTGAGAAACAGAGAGTGGCACCGCCAGAGATCCCTTTTA 1683
 Qy 1501 TTTCCATGCAATGCAATGAGGAGGTGAGAGACAAATTACACACAGCTTCTTTTGGC 1560
 Db 1684 TTTCCATGCAATGCAATGAGGAGGTGAGAGACAAATTACACACAGCTTCTTTTGGC 1743
 Qy 1561 CTTTCCACGAGCCAGAGACCTTCAAGAAATGTGCTGGCTGGGCTTAAAGGCTGGCAC 1620
 Db 1744 CTTTCCACGAGCCAGAGACCTTCAAGAAATGTGCTGGCTGGGCTTAAAGGCTGGCAC 1803
 Qy 1621 TGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGAGCTGATTTTGGC 1680
 Db 1804 TGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGAGCTGATTTTGGC 1863
 Qy 1681 ACAAGTCTTCACTTTACTCTGCTGAGCATCTTACCGAGAGCAGTGCAGTACTCT 1740
 Db 1864 ACAAGTCTTCACTTTACTCTGCTGAGCATCTTACCGAGAGCAGTGCAGTACTCT 1923
 Qy 1741 GCTTACAGCTGACAGCAGCTGCCACTTGCAGAGACCAAGTATTTGTGCGACGCG 1800
 Db 1924 GCTTACAGCTGACAGCAGCTGCCACTTGCAGAGACCAAGTATTTGTGCGACGCG 1983
 Qy 1801 CAGAAAGCAAGTACAGAGCTGCTCCGCGGAGCTGGCATGAAGAGGCCCTTTGA 2043
 Db 1984 CAGAAAGCAAGTACAGAGCTGCTCCGCGGAGCTGGCATGAAGAGGCCCTTTGA 2103
 Qy 1861 AAGCAGTTTAAACCAAGAGCTGCAAAATGGAATTTGAGAGAGATCATGTCAGAGAAC 1920
 Db 2044 AAGCAGTTTAAACCAAGAGCTGCAAAATGGAATTTGAGAGAGATCATGTCAGAGAAC 2103
 Qy 1921 AGTTCACGGGAAAGCTGGGGAAAGTGGGCACTGCTTTTGGGCGAGATGGAA 1980
 Db 2104 AGTTCACGGGAAAGCTGGGGAAAGTGGGCACTGCTTTTGGGCGAGATGGAA 2163
 Qy 1981 ATCATTTAGGTTCTCC 1995
 Db 2164 ATCATTTAGGTTCTCC 2178

RESULT 9
 ABV20833
 ID ABV20833 standard; cDNA; 5145 BP.
 AC ABV20833;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 20824.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX
 PN W0200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001MO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183319P.

Db 2209 TCGATATCTTGGCCCCCGAGCCTTACCCCTTCCCTGACAGACAGCTGTATTTTGGC
Qy 1681 ACAGAGTCTTCACTTCTTACTTCTGCTCAGCCATCTACGAGAGAGTGCAGTTACTCT 2268
Db 2269 AAGAGTCTTCACTTCTTACTTCTGCTCAGCCATCTACGAGAGAGTGCAGTTACTCT 1740
Qy 1741 GCTTACAGCTGAGCAGCTGCTCCCACTTGGGAGAGCCAACTGTATTTCTGCGAGAGCG 2328
Db 2329 GCTTACAGCTGAGCAGCTGCTCCCACTTGGGAGAGCCAACTGTATTTCTGCGAGAGCG 1800
Qy 1801 CAGAGCCAACTGTACAGAGCTGACTGCGCGGAGCTGAGATGAGAGAGAGCCCTTTGAA 2388
Db 2389 CAGAGCCAACTGTACAGAGCTGACTGCGCGGAGCTGAGATGAGAGAGAGCCCTTTGAA 1860
Qy 1861 AAGCAATTTTAAACGAGAGCTGCTCCAAATGGAATTTGAGAGAGCATGTACAGAGAC 2448
Db 2449 AAGCAATTTTAAACGAGAGCTGCTCCAAATGGAATTTGAGAGAGCATGTACAGAGAC 1920
Qy 1921 AGGTACAGGAGAGAGCTGCGGAGAGTGGCACTAGCTTTTGGGAGAGATGGA 2508
Db 2509 AGGTACAGGAGAGAGCTGCGGAGAGTGGCACTAGCTTTTGGGAGAGATGGA 1980
Qy 1981 ATCATTTAGAGTCTCC 1995
Db 2569 ATCATTTAGAGTCTCC 2583

RESULT 10

ABV21080
ID ABV21080 standard, cDNA, 5145 BP.

XX ABV21080;

DT 13-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 21071.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN W0200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US005171.

PR 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JB;

XX WPI; 2001-662795/76.

DR Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
PS Claim 1; Page 3481; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 99.8%; Score 1991.8; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGCCCCATGATGATGATGGAACCTAAATGTTACTGAGAGGTGGTGGCTGCTGGA 60
Db 589 ATGCCCCATGATGATGATGGAACCTAAATGTTACTGAGAGGTGGTGGCTGCTGGA 648
Qy 61 AGTGAACGGAAGAGAGCTGCTTAAATGATAGCCGCGCATTTGTGAAATACATATATCC 120
Db 649 AGTGAACGGAAGAGAGCTGCTTAAATGATAGCCGCGCATTTGTGAAATACATATCC 708
Qy 121 CACATTTTGAAGCCATTAATATCACTGCTCAAGCTTATGAGAGAGGTGCAACAG 180
Db 709 CACATTTTGAAGCCATTAATATCACTGCTCAAGCTTATGAGAGAGGTGCAACAG 768
Qy 181 GACAAAGTGTATATACAGAGCTCAATCCAGATTCAGAGAAATTAAGTTGATTTGAT 240
Db 769 GACAAAGTGTATATACAGAGCTCAATCCAGATTCAGAGAAATTAAGTTGATTTGAT 828
Qy 241 TGCAGTCAAGAGTGTATGATTACAGTCAAGCTCCAGAGATTTGCTCTCTCTCA 300
Db 829 TGCAGTCAAGAGTGTATGATTACAGTCAAGCTCCAGAGATTTGCTCTCTCTCA 888
Qy 301 GACTGTTTCTCACTGATCTCTGAGTAACTGGAAGAGCTTCAACTGTGTCACCTG 360
Db 889 GACTGTTTCTCACTGATCTCTGAGTAACTGGAAGAGCTTCAACTGTGTCACCTG 948
Qy 361 CTGAGAGTGGGTTTCTGAGTCTCTGTTGTTTCCCTGAGCTGTAAGAGAAATTC 420
Db 949 CTGAGAGTGGGTTTCTGAGTCTCTGTTGTTTCCCTGAGCTGTAAGAGAAATTC 1008
Qy 421 ACTTAGTCCCTACCTGATTTCTCAGCCCTGTTACCTGTTGCCAATCTGGCCAAAC 480
Db 1009 ACTTAGTCCCTACCTGATTTCTCAGCCCTGTTACCTGTTGCCAATCTGGCCAAAC 1068
Qy 481 GGAATTCCTCCCAATCTTATCTTGGCTCCAGAGATGCTCTCAACAGAGCTGATA 540
Db 1069 GGAATTCCTCCCAATCTTATCTTGGCTCCAGAGATGCTCTCAACAGAGCTGATA 1128
Qy 541 CAGCAGATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 1129 CAGCAGATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGATGATGATGATGATGATGATGATGATGAT 1188
Db 1189 ATCCCGAGTCTCATTTCTGCGTGTGCTGATGATGATGATGATGATGATGATGATGAT 1248
Qy 661 CCGTGTGGAACAAATCAGTATGATTTCTTGAAGAAACAAAGCTTCATGATGATGAT 720
Db 1249 CCGTGTGGAACAAATCAGTATGATTTCTTGAAGAAACAAAGCTTCATGATGATGAT 1308
Qy 721 CTAGTGAATGTTTATGCTGGAGATCTCCGCTCCGCAACATGCTTATGCTATGATGAT 780
Db 1309 CTAGTGAATGTTTATGCTGGAGATCTCCGCTCCGCAACATGCTTATGCTATGATGAT 1368
Qy 781 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 1369 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Qy 841 ATATCTCCAAATCTTATTTCTGGGCAATCTGAGATGATGATGATGATGATGATGATGAT 1428
Db 1429 ATATCTCCAAATCTTATTTCTGGGCAATCTGAGATGATGATGATGATGATGATGATGAT 1488

Qy	901	AAGCTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCTGACCTGAGAAAGCAAT	960
Db	1489	CAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCTGACCTGAGAAAGCAAT	15448
Qy	961	GAACTGTCTCCGTGTCTCAGAGGGTGAACAGAAAGCGAGACGCCCTCCAGTCCACC	1020
Db	1549	GAACTGTCTCCGTGTCTCAGAGGGTGAACAGAAAGCGAGACGCCCTCCAGTCCACC	1608
Qy	1021	TGTGCGGACTGTGCTACCTCAGAGGACACAGAGCAAAAGCCCGTGTCAATCCCGCAGCGTG	108
Db	1609	TGTGCGGACTGTGCTACCTCAGAGGACACAGAGCAAAAGCCCGTGTCAATCCCGCAGCGTG	1668
Qy	1081	CCCAAGCTGTCCCAAGCGTGCAGCCGTGCGCTGTGAGAGACAGCCGCGCTGTACAGAGCGTGC	1144
Db	1669	CCCAAGCTGTCCCAAGCGTGCAGCCGTGCGCTGTGAGAGACAGCCGCGCTGTACAGAGCGTGC	1728
Qy	1141	AGTGGGCTGCACCTGTCCGCAGACAGGCTGGAAGACAGCAATAGCTCAAGCTTCTTTC	1200
Db	1729	AGTGGGCTGCACCTGTCCGCAGACAGGCTGGAAGACAGCAATAGCTCAAGCTTCTTTC	1788
Qy	1201	TCTCTGGATATCAATCAATTCAATATTCAGCGACATGGAGCATCTCTTCAATGGCTTC	126
Db	1789	TCTCTGGATATCAATCAATTCAATATTCAGCGACATGGAGCATCTCTTCAATGGCTTC	1848
Qy	1261	TCCTCATCAGAAAGATGCTTTGGAAATACTACAAACCTTCCACTACTCTGGATGGACCAAC	1320
Db	1849	TCCTCATCAGAAAGATGCTTTGGAAATACTACAAACCTTCCACTACTCTGGATGGACCAAC	1908
Qy	1321	AAGCTATGCCAGTTTCTCCCTGTTCAGAGACTATTTGGAGCAGACTCCGAAAACCACTCT	1380
Db	1909	AAGCTATGCCAGTTTCTCCCTGTTCAGAGACTATTTGGAGCAGACTCCGAAAACCACTCT	1968
Qy	1381	GATTAAGAGAGAACCCAGCATCCCAACAAAGCTGCAGACCGCACGAGCTTCAGACAGCCAG	1440
Db	1969	GATTAAGAGAGAACCCAGCATCCCAACAAAGCTGCAGACCGCACGAGCTTCAGACAGCCAG	2028
Qy	1441	AGCAAGCGATTGCAATTGCTGCTAGAAACAGCAGCAGTGGCACCGCCAGAGGTCCCTTTTA	1500
Db	2029	AGCAAGCGATTGCAATTGCTGCTAGAAACAGCAGCAGTGGCACCGCCAGAGGTCCCTTTTA	2088
Qy	1501	TCTCCACATCGATCCAAAGTGGGAGGCTGGAGGACAAATTACACACAGCTTCTTTTGGG	1566
Db	2089	TCTCCACATCGATCCAAAGTGGGAGGCTGGAGGACAAATTACACACAGCTTCTTTTGGG	2148
Qy	1561	CTTTTCCACAGACCGACAGCACCTCCACGAAGTCTGCTGGCTTGGGCTTTAAGGGCTGGCAC	1620
Db	2149	CTTTTCCACAGACCGACAGCACCTCCACGAAGTCTGCTGGCTTGGGCTTTAAGGGCTGGCAC	2208
Qy	1621	TCCGATATCTTTGGCCCCCAGACCTCTAACCCCTTCCCTGACACAGCATGTGTATTTTGGC	1680
Db	2209	TCCGATATCTTTGGCCCCCAGACCTCTAACCCCTTCCCTGACACAGCATGTGTATTTTGGC	2268
Qy	1681	ACAGAGTCTTCACACTTCTACCTGCGCCACAGCAATCTACGAGAGGACAGTGCATTACTCT	1740
Db	2269	ACAGAGTCTTCACACTTCTACCTGCGCCACAGCAATCTACGAGAGGACAGTGCATTACTCT	2328
Qy	1741	GCCCTACAGCTGCAGCCAGCTGCCCATTTGCCGAGACCAAGTCTATTTCTGTGCGCAGCGG	1800
Db	2329	GCCCTACAGCTGCAGCCAGCTGCCCATTTGCCGAGACCAAGTCTATTTCTGTGCGCAGCGG	2388
Qy	1801	CAGAAAGCAATGACAGAGCTGACCTCCGGGGGGAGCTGGCATTGAAGAAAGCCCTTTGAA	1860
Db	2389	CAGAAAGCAATGACAGAGCTGACCTCCGGGGGGAGCTGGCATTGAAGAAAGCCCTTTGAA	2448
Qy	1861	AAGCAGTTTAAACCGCAGAGCTGCCCAATTGGAAATTTGGAGAGAGCATCATGTCAAGAAC	1920
Db	2449	AAGCAGTTTAAACCGCAGAGAGTGCACCAATTGGAAATTTGGAGAGAGCATCATGTCAAGAAC	2508
Qy	1921	AGGTTCACGGGAAAGAGCTGGGGAAAGTGGGCACTCAGTCACTTTTCTGGGCGACATGGAA	1980
Db	2509	AGGTTCACGGGAAAGAGCTGGGGAAAGTGGGCACTCAGTCACTTTTCTGGGCGACATGGAA	2568

Y	1981	ATCATTTGAGTCTCC	1995
Db	2569	ATCATTTGAGTCTCC	2583
	RESULT 11		
	ABV26680		
	ABV26680	standard; cDNA; 5145 BP.	
XX	AC	ABV26680;	
XX	DT	16-SEP-2002 (first entry)	
XX	DE	Human prostate expression marker cDNA 26671.	
XX	KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;	
XX	KW	pharmacogenomic marker; gene; ss.	
XX	OS	Homo sapiens.	
XX	PN	WO200160860-A2.	
XX	PD	23-AUG-2001.	
XX	PF	20-FEB-2001; 2001WO-US005171.	
XX	PR	17-FEB-2000; 2000US-0183319P.	
XX	PR	16-MAR-2000; 2000US-0189862P.	
XX	PR	25-MAY-2000; 2000US-0207454P.	
XX	PR	09-JUN-2000; 2000US-0211314P.	
XX	PR	18-JUL-2000; 2000US-0219007P.	
XX	PR	13-DEC-2000; 2000US-0255281P.	
XX	PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.	
XX	PI	Schlegel R, Endege WO, Monahan JE;	
XX	DR	WPI; 2001-662795/76.	
XX	PT	Novel isolated nucleic acid molecule associated with cancerous state of	
XX	PT	prostate cells and correlating with presence of prostate cancer, useful	
XX	XX	for detecting presence of prostate cancer, stage of prostate cancer.	
XX	PS	Claim 1; Page 5388-5389; 11750pp; English.	
XX	CC	The invention relates to an isolated nucleic acid molecule (I) comprising	
XX	CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the	
XX	CC	specification or its complement. (I) is useful for: (a) assessing whether	
XX	CC	a patient is afflicted with prostate cancer; (b) monitoring the	
XX	CC	progression of prostate cancer in a patient; (c) assessing the efficacy	
XX	CC	of a test compound to inhibit prostate cancer in a patient; (d) assessing	
XX	CC	the efficacy of a therapy for inhibiting prostate cancer in a patient;	
XX	CC	(e) selecting a composition for inhibiting prostate cancer in a patient;	
XX	CC	(f) assessing the prostate cell carcinogenic potential of a compound; (g)	
XX	CC	determining whether prostate cancer has metastasized in a patient; (h)	
XX	CC	assessing the aggressiveness or indolence of prostate cancer in a patient	
XX	CC	; (I) is also useful as a pharmacodynamic or pharmacogenomic marker	
XX	XX	Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;	
QY	Query Match	99.8%; Score 1991.8; DB 5; Length 5145;	
Db	Best Local Similarity	99.9%; Pred. No. 0;	
	Matches 1993; Conservative	0; Mismatches 2; Indels 0; Gaps 0	
QY	1	ATGSCCATGATGATGATGGAATCTCAATTGTAATGAGAGCTTGCTGCTGGAA	60
Db	589	ATGSCCATGATGATGATGGAATCTCAATTGTAATGAGAGCTTGCTGCTGGAA	648
QY	61	AGTGAACGGAAGAGTCTCTAATTGATATGAGCGGCATTTGTGGAATACATATACCC	120
Db	649	AGTGAACGGAAGAGTCTCTAATTGATATGAGCGGCATTTGTGGAATACATATACCC	708
QY	121	CACATTTTGGAAGCATTAATATCACTGCTCAAGCTTATGAAGCAAGGTTCAACAG	180

XX 20-FEB-2001; 2001WO-US005171.
XX
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0235281P.
XX
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1, Page 3451, 11750pp, English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (c) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (d) selecting a composition for inhibiting prostate cancer in a patient;
CC (e) assessing the prostate cell carcinogenic potential of a compound; (f)
CC determining whether prostate cancer has metastasized in a patient; (g)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
Query Match 99.8%; Score 1991.8; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCATGAGATGATTTGGAACCTCAATTTGTTACTGAGAGTTGGTGGCTGTGAGAA 60
DB 589 ATGGCCATGAGATGATTTGGAACCTCAATTTGTTACTGAGAGTTGGTGGCTGTGAGAA 648
QY 61 AGTGAACGGAAGAGTGTCTGTAATGATAGCCGGCATTGTTGGAATCAATATACATCC 120
DB 649 AGTGAACGGAAGAGTGTCTGTAATGATAGCCGGCATTGTTGGAATCAATATACATCC 708
QY 121 CAAATTTGGAACCTCAATTTGGAACCTCAATTTGGAACCTCAATTTGGAACCTCAATTTG 180
DB 709 CAAATTTGGAACCTCAATTTGGAACCTCAATTTGGAACCTCAATTTGGAACCTCAATTTG 768
QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATTGAT 240
DB 769 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATTGAT 828
QY 241 TGCAGTCAGAGGTTGATTTACATCAAGAGCTCCCAAGATGTTGCTCTCTCTTCA 300
DB 829 TGCAGTCAGAGGTTGATTTACATCAAGAGCTCCCAAGATGTTGCTCTCTCTTCA 888
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACCTGAGAGAGCTTCAACTGTTTCACTG 360
DB 889 GACTGTTTCTCACTGATCTTCTGGGTAACCTGAGAGAGCTTCAACTGTTTCACTG 948
QY 361 CTTCAGAGTGGTGTGAGTCTCTCGTTGTTCCCTGGCTCTGTGAGAGAAATCC 420
DB 949 CTTCAGAGTGGTGTGAGTCTCTCGTTGTTCCCTGGCTCTGTGAGAGAAATCC 1008
QY 421 ACTCTAGTCCCTTACCTGATTTCTCAGCCTTGTACCTGTTGCTTCAACATTTGGGCAAC 480
DB 1009 ACTCTAGTCCCTTACCTGATTTCTCAGCCTTGTACCTGTTGCTTCAACATTTGGGCAAC 1068

QY 481 CGAATTCCTCCCAATCTTTATCTTGGCTGCAGAGATGTCCTCAACAGAGCTGATA 540
DB 1069 CGAATTCCTCCCAATCTTTATCTTGGCTGCAGAGATGTCCTCAACAGAGCTGATA 1128
QY 541 CACAGAGATGGATGTTGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 600
DB 1129 CACAGAGATGGATGTTGTTATGTTTAAATGCCAGGAAATACCTGTCCAAAGCTGACTTT 1188
QY 601 ATCCCGAGTCTATTTCTGCTGCTGTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG 660
DB 1189 ATCCCGAGTCTATTTCTGCTGCTGTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG 1248
QY 661 CCGTGGTGGACAAATCAAGATTTTCAATGAGAAAGCAAAAGCTCCAAATGATGTT 720
DB 1249 CCGTGGTGGACAAATCAAGATTTTCAATGAGAAAGCAAAAGCTCCAAATGATGTT 1308
QY 721 CTATGCACTGTTTAACTGGGATCTCCGCTCCGCAACATTCCTATGCTTATCATCATG 780
DB 1309 CTATGCACTGTTTAACTGGGATCTCCGCTCCGCAACATTCCTATGCTTATCATCATG 1368
QY 781 AAGAGATGGAATGCTTCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGACTACT 840
DB 1369 AAGAGATGGAATGCTTCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGACTACT 1428
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGCACTATGAGAAAGATTAAGAAC 900
DB 1429 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGCACTATGAGAAAGATTAAGAAC 1488
QY 901 CAGACTGAGCATCAGAGGCCAAAGAGCAAACTCAAGCTGTGCACTGAGAGAGCCAAAT 960
DB 1489 CAGACTGAGCATCAGAGGCCAAAGAGCAAACTCAAGCTGTGCACTGAGAGAGCCAAAT 1548
QY 961 GAACTGCTGCTGCTGTCTCAAGAGGTGACAGAAAAGGAGAGCCCTCACTCAACC 1020
DB 1549 GAACTGCTGCTGCTGTCTCAAGAGGTGACAGAAAAGGAGAGCCCTCACTCAACC 1608
QY 1021 TGTGCGGACCTGCTACCTCAAGAGGAGAGCAAAAGCCGCTGATCCGCGAGGTG 1080
DB 1609 TGTGCGGACCTGCTACCTCAAGAGGAGAGCAAAAGCCGCTGATCCGCGAGGTG 1668
QY 1081 CCCAGGCTGCCAGCTGAGCCGCTGCTGTTAGAGAGACAGCCGCTGATCAAGGCTC 1140
DB 1669 CCCAGGCTGCCAGGCTGAGCCGCTGCTGTTAGAGAGACAGCCGCTGATCAAGGCTC 1228
QY 1141 AGTGGCTGCACTGCTGCTGCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
DB 1729 AGTGGCTGCACTGCTGCTGCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1788
QY 1201 TCTCTGATATCAATCAATTTCAATTTCAAGGAGATGGAGCATCTTACATGCTTC 1260
DB 1789 TCTCTGATATCAATCAATTTCAATTTCAAGGAGATGGAGCATCTTACATGCTTC 1848
QY 1261 TCTCTATCAGAAAGATGCTTGGATATCAATCAATTTCAATTTCAATTTCAATTTCA 1320
DB 1849 TCTCTATCAGAAAGATGCTTGGATATCAATCAATTTCAATTTCAATTTCAATTTCA 1908
QY 1321 AAGCTATGCCAGTCTTCCCTGTTCAAGAACTATGAGAGCACTCCCGAACAAGTCTC 1380
DB 1909 AAGCTATGCCAGTCTTCCCTGTTCAAGAACTATGAGAGCACTCCCGAACAAGTCTC 1968
QY 1381 GATTAAGAGAGAGGAGGAGATCCCAAGAGCTGAGACAGGCGCTTCAAGACAGCAG 1440
DB 1969 GATTAAGAGAGAGGAGGAGATCCCAAGAGCTGAGACAGGCGCTTCAAGACAGCAG 2028
QY 1441 AGCAAGGATTCATTCGCTGAGAAACCAAGAGAGTGGACAGCCGCAAGAGTCTCTTTTA 1500
DB 2029 AGCAAGGATTCATTCGCTGAGAAACCAAGAGAGTGGACAGCCGCAAGAGTCTCTTTTA 2088
QY 1501 TCTTCATCTGATGAAAGTGGAGGCTGAGAGCAATTAACACACAGCTTCTTTTGGG 1560
DB 2089 TCTTCATCTGATGAAAGTGGAGGCTGAGAGCAATTAACACACAGCTTCTTTTGGG 2148
QY 1561 CTTTCAACAGCAGAGAGACCTCAAGAACTGCTGCTGGGCTTAAAGGCTGGCAC 1620

Db 2149 CTTTCACACACCCAGCAGCACTTCAGAAAGTCTGCTGGCCCTTGAAGGCTGGCAC 2208
QY 1621 TCGGATATCTTGGGCCCCCGAGACCTTCAACCCCTTCCCTGACACAGACCTGTATATTTGGC 1681
Db 2209 TCGGATATCTTGGGCCCCCGAGACCTTCAACCCCTTCCCTGACACAGACCTGTATATTTGGC 2268
QY 1681 ACAGAGTCTTCACACCTTTTACTCTTCGCTCAGCACTTACAGGAGGAGTGCAGTTACTCT 1740
Db 2269 ACAGAGTCTTCACACCTTTTACTCTTCGCTCAGCACTTACAGGAGGAGTGCAGTTACTCT 2328
QY 1741 GCCTTACAGCTGCAGCGCAGCTGGCCACTTGGCGGAAACAAGTCTATTTCTGTGGCAGGCGG 1800
Db 2329 GCCTTACAGCTGCAGCGCAGCTGGCCACTTGGCGGAAACAAGTCTATTTCTGTGGCAGGCGG 2388
QY 1801 CAGAACCCAAAGTGAAGAAGCTGACTCGCGGCGGAGCTGTGATTAAGAGAACCCCTTTGAA 1860
Db 2389 CAGAACCCAAAGTGAAGAAGCTGACTCGCGGCGGAGCTGTGATTAAGAGAACCCCTTTGAA 2448
QY 1861 AAGCAGTTTAAACGCAAGAGCTGCCAATGGAATTTGAGAGAGCAGTATCTCAGAGAC 1920
Db 2449 AAGCAGTTTAAACGCAAGAGCTGCCAATGGAATTTGAGAGAGCAGTATCTCAGAGAC 2508
QY 1921 AGCTCACCGGGAAGAGCTGTGGGAAAGTGGCAGTCAGCTAGCTTTTGGGCGAGCATGGA 1980
Db 2509 AGCTCACCGGGAAGAGCTGTGGGAAAGTGGCAGTCAGCTAGCTTTTGGGCGAGCATGGA 2568
QY 1981 ATCATTTAGGAGCTCC 1995
Db 2569 ATCATTTAGGAGCTCC 2583

RESULT	13
ABV21092	
ID	ABV21092 standard; cDNA; 5145 BP.
XX	

XX	13-SEP-2002	(first entry)
DT		
XX		

Human prostate expression marker cDNA 21083.

pharmacogenomic marker; gene; ss.

Homo sapiens.

WC200160860-A2.

23-AUG-2001.

20-FEB-2001; 2001WO-US005171.

16-MAR-2000; 2000US-0189863P

09-JUN-2000; 2000US-0217314P

13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schreyer R, Endege WO, Monahan JE;

96/561799-T007, T008

Claim 1, Page 3485, 11750bp, English

the invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or incidence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
CC
CC Sequence 5145 BP, 1408 A; 1135 C; 1253 G; 1146 T; 0 nt; 1 nt; 2 nt;
CC

Query Match	99.8%	Score 1991.8;	DB 5;	Length 5145;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1993; Conservative	0;	Mismatches	2;	Indels 0.

[illegible]

649 AATTGGAAAGGAAAAAGTGTCTTAATTGATAGCCGGCAATTGGAAATCAATTCATCC 708
 121 CACATTTTGGAGCATTAATATCAAGTCGTCACGCTTTGAAGGAAGGTTGCACAG 180

[illegible]

241 TGCAGTCAGAAAGGTGTAGTTTACGATCAAAAGCTCCCAAGATGTTGCCTCTCTCTTCA 828
 |||||
 829 TGCAGTCAGAAAGGTGTAGTTTACGATCAAAAGCTCCCAAGATGTTGCCTCTCTCTTCA 300
 |||||
 301

361 CTTGACGGTGGTTTGTGCTGAACTCTCTGTTGTTTCCCTGGACCTCTGAAAGGAAATATTC 420

[illegible]

481 CGAATTCCTCCAACTCTTATCTTGGCTGCCAGAGATGCTCCCAACAAGAGCTGTATA 540
|||||
1069 CGAATTCCTCCAACTCTTATCTTGGCTGCCAGAGATGCTCCCAACAAGAGCTGTATG 1128
|||||

541 CAGCGAATGGGANTGGTATGTCTTAATGCCACGTATCCGTGTCAGAAGCCTGACTTT 600
1129 CAGCAAGATGGGATTGGTATGTCTTAATGCCACGAATCTCGTGCAGAAAGCTGCACTTT 1188
601 ATCCCGAGTCTCATTTCTCGGTGTGTCCTGCTCAATAATGACCTGCTGCTGCTGCTGCT

1189 ATCCCCAGTCTATTTCCTGCTGTGCTGTGAATGACAGCTTTTGACAAAATTTTG 1248
|||||
661 CCGGTGTGGACAAATCAGTGAATTCATTGAGAAAGCAAAAGCCTCCATGCAATGTGTT 720
|||||

721 CTATGCACTGTTTAGCGGGGATTCCTCCGTCACCATGCTATGCGCTCAATCATG 780
1309 CTATGCACTGTTTAGCGGGGATTCCTCCGTCACCATGCTATGCGCTCAATCATG 1360

781 AAGAGATGACATCTTTAGATGAAGCTTACAGTTTGTGAAGAAAAAAGACTTCT 840
1369 AAGAGATGACACATCTTTTAGATGAAGCTTACAGATTGTGAAGAAAAAAGACTTCT 1428

QY 841 ATATTCCTCAAACTTCAATTTTCTGGGCGCACTCCTGSACTATGAGAGATTTAAGAAC 900
DB 1429 ATATCTCCAACTTCAATTTTCTGGGCGCACTCCTGSACTATGAGAGATTTAAGAAC 1488
QY 901 CAGACTGAGCATCAGGGCCAAAGAGAACTCAAGCTGTGCACTTGAGAGCCCAAT 960
DB 1489 CAGACTGAGCATCAGGGCCAAAGAGAACTCAAGCTGTGCACTTGAGAGCCCAAT 1548
QY 961 GAACCTGTCCCTGTCTCTCAAGAGGTGACAGAAAAAGAGAGCCCTCAGTCCACC 1020
DB 1549 GAACCTGTCCCTGTCTCTCAAGAGGTGACAGAAAAAGAGAGCCCTCAGTCCACC 1608
QY 1021 TGTGCGCACTGTCTCTCAAGAGGTGACAGAAAAAGAGAGCCCTCAGTCCACC 1080
DB 1609 TGTGCGCACTGTCTCTCAAGAGGTGACAGAAAAAGAGAGCCCTCAGTCCACC 1668
QY 1081 CCCAGCGTCCAGCGCTGACAGCGCTGCTGTAGAGAGACGCCGCTGTACAGGCGCTC 1140
DB 1669 CCCAGCGTCCAGCGCTGACAGCGCTGCTGTAGAGAGACGCCGCTGTACAGGCGCTC 1728
QY 1141 AGTGGGCTGACCTGTCTCGAGACAGGCTGAAAGACAGCAATAGCTCAAGCGTCTTC 1200
DB 1729 AGTGGGCTGACCTGTCTCGAGACAGGCTGAAAGACAGCAATAGCTCAAGCGTCTTC 1788
QY 1201 TCTCTGATATCAAAATGATTTTCAATTTCAAGCCAGCATGGCAGATCTTACATGGCTTC 1260
DB 1789 TCTCTGATATCAAAATGATTTTCAATTTCAAGCCAGCATGGCAGATCTTACATGGCTTC 1848
QY 1261 TCTCTATCAAGAAATGCTTTGSAATCTAATAAATCTTCACTCTGTAGATGGAGCCAA 1320
DB 1849 TCTCTATCAAGAAATGCTTTGSAATCTAATAAATCTTCACTCTGTAGATGGAGCCAA 1308
QY 1321 AAGCTATGCAAGTTCTCCCTGTTGAGAACTATCGAGAGACCTCCGAAACCAAGTCT 1380
DB 1909 AAGCTATGCAAGTTCTCCCTGTTGAGAACTATCGAGAGACCTCCGAAACCAAGTCT 1968
QY 1381 GATTAAGAGAGAACCCAGCATCCCAAGAACTGACAGCCCGAGGCTTTAGACAGCCAG 1440
DB 1969 GATTAAGAGAGAACCCAGCATCCCAAGAACTGACAGCCCGAGGCTTTAGACAGCCAG 2028
QY 1441 AGCAAGGATTTGATTCGTCAGAACAGAGAGAGTGGAGCCAGGCTTCCCTTTTA 1500
DB 2029 AGCAAGGATTTGATTCGTCAGAACAGAGAGAGTGGAGCCAGGCTTCCCTTTTA 2088
QY 1501 TCTCCACTGATCGAAGTGGAGAGGTGAGAGACAAATTACCAACAGACTTCTTTTGCGC 1560
DB 2089 TCTCCACTGATCGAAGTGGAGAGGTGAGAGACAAATTACCAACAGACTTCTTTTGCGC 2148
QY 1561 CTTTCCACCCAGCCAGAGCACTTCAGAAATGCTGTGGCTTGGGCTTTAAGGCTGGCAC 1620
DB 2149 CTTTCCACCCAGCCAGAGCACTTCAGAAATGCTGTGGCTTGGGCTTTAAGGCTGGCAC 2208
QY 1621 TCGGAATTCCTGGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTATTTTGGC 1680
DB 2209 TCGGAATTCCTGGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTATTTTGGC 2268
QY 1681 ACAGAGTCTCAGACTTCTAATCTGTGCTCAGCACTTACGAGAGAGAGTGCAGTTACTCT 1740
DB 2269 ACAGAGTCTCAGACTTCTAATCTGTGCTCAGCACTTACGAGAGAGAGTGCAGTTACTCT 2328
QY 1741 GCTTACAGCTGACCCAGAGCTGCCCACTTGGGAGACCTAATCTAATTTGTGGCCAGGCGG 1800
DB 2329 GCTTACAGCTGACCCAGAGCTGCCCACTTGGGAGACCTAATCTAATTTGTGGCCAGGCGG 2388
QY 1801 CAGAAAGCAAGTGCAGAGCTGACTCGGCGAGAGTGCATGATGAAGAGAGCCCTTTGAA 1860
DB 2389 CAGAAAGCAAGTGCAGAGCTGACTCGGCGAGAGTGCATGATGAAGAGAGCCCTTTGAA 2448
QY 1861 AAGCAGTTTAAACGACAGAGCTGCAAAATGAAATTTTGAAGAGAGCATGTGACAGAGAAC 1920
DB 2449 AAGCAGTTTAAACGACAGAGCTGCAAAATGAAATTTTGAAGAGAGCATGTGACAGAGAAC 2508

QY 1921 AGGTACGGGAGAGAGCTGGGAGAAAGTGGGAGTCACTAGCTTTTGGGACATGGA 1980
DB 2509 AGGTACGGGAGAGAGCTGGGAGAAAGTGGGAGTCACTAGCTTTTGGGACATGGA 2568
QY 1981 ATCATTGAGGTCTCC 1995
DB 2569 ATCATTGAGGTCTCC 2583

RESULT 14
ABV21312
ID ABV21312 standard; cDNA; 5145 BP.
XX
AC ABV21312;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 21303.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN W02001.60860-A2.
XX
PD 23-ANG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-018319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3539; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 99.8%; Score 1991.8; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCGCCATGAGATGATTTGGAAGTCAATTTGTTACTGAGAGAGTGGTGTCTGCTGGAA 60
DB 589 ATGCGCCATGAGATGATTTGGAAGTCAATTTGTTACTGAGAGAGTGGTGTCTGCTGGAA 648
QY 61 AGTGAAGCGAAGAAAGTGTGCTAATTTGATTTAGCCGCGCAATTTGTGAAATACATATCCTC 120

Db 649 AGTGAAGGAAAAAGTCTCTCTAATTGATAGCGGCGCATTTGTGAAATACATATCATCC 708
 QY 121 CACATTTTGAAGCATTAATATATCACTGCTCCAGGCTTATGAAGGAGTTGCAACAG 180
 Db 709 CACATTTTGAAGCATTAATATCACTGCTCCAGGCTTATGAAGGAGTTGCAACAG 768
 QY 181 GACAAAGTTAATTAAGAGCTCATCAGCATTCAGGAACTTAAGTTGACATTGAT 240
 Db 769 GACAAAGTTAATTAAGAGCTCATCAGCATTCAGGAACTTAAGTTGACATTGAT 828
 QY 241 TGCAGTGAAGGTTGATTAATTAAGATCAAGCTCCAGAGATGTTGCTCTCTCTTCA 300
 Db 829 TGCAGTGAAGGTTGATTAATTAAGATCAAGCTCCAGAGATGTTGCTCTCTCTTCA 888
 QY 301 GACTGTTTCTCATCTGACTTCTGAGGTTAACTGGAAGAGTTTAACTCTGTTACCTG 360
 Db 889 GACTGTTTCTCATCTGACTTCTGAGGTTAACTGGAAGAGTTTAACTCTGTTACCTG 948
 QY 361 CTGCAAGGTTGCTGAGTTCTCTGTTGTTTCCCTGCGCTCTGTGAAGAAATCC 420
 Db 949 CTGCAAGGTTGCTGAGTTCTCTGTTGTTTCCCTGCGCTCTGTGAAGAAATCC 1008
 QY 421 ACTAGTCCCTACCTGACTTCTCAGGCTTGTACTTCTGTCGCAACTTGGGCAACC 480
 Db 1009 ACTAGTCCCTACCTGACTTCTCAGGCTTGTACTTCTGTCGCAACTTGGGCAACC 1068
 QY 481 CGAATTTCTCCCAATCTTTATCTTGTGCTGCGCAGGAGATGTCCTCAAGAGCTGATG 1128
 Db 1069 CGAATTTCTCCCAATCTTTATCTTGTGCTGCGCAGGAGATGTCCTCAAGAGCTGATG 1128
 QY 541 CAGCAGAAATGGGATGTTATGTTTAAATGCGGCTTATCTGTCCAAAGCTGACTT 600
 Db 1129 CAGCAGAAATGGGATGTTATGTTTAAATGCGGCTTATCTGTCCAAAGCTGACTT 1188
 QY 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGAAGCTTTTGTGAAGAAATTTTG 660
 Db 1189 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGAAGCTTTTGTGAAGAAATTTTG 1248
 QY 661 CCGTGTGTGAAGAAATCAGTAGATTTCAATTGAGAAGCAAAAGCTTCCAAATGATGTT 720
 Db 1249 CCGTGTGTGAAGAAATCAGTAGATTTCAATTGAGAAGCAAAAGCTTCCAAATGATGTT 1308
 QY 721 CTAGTGAATGTTAGTGTGAGATCTCCGCTCCGCAACCATGCTATGCTTACATCATG 780
 Db 1309 CTAGTGAATGTTAGTGTGAGATCTCCGCTCCGCAACCATGCTATGCTTACATCATG 1368
 QY 781 AAGAGATGAGCATGTCTTTAAGTAGAAGCTTACAGATTTTGAAGAAAGAAAGCTTACT 840
 Db 1369 AAGAGATGAGCATGTCTTTAAGTAGAAGCTTACAGATTTTGAAGAAAGAAAGCTTACT 1428
 QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAGAAAGATTAGAAC 900
 Db 1429 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAGAAAGATTAGAAC 1488
 QY 901 CAGACTGAGCATCAGGCGCAAAAGAGCAAACTCAAGCTGTGCACTTGTGAAGAACCAAT 960
 Db 1489 CAGACTGAGCATCAGGCGCAAAAGAGCAAACTCAAGCTGTGCACTTGTGAAGAACCAAT 1548
 QY 961 GAACCTGTCCCTGCTGTCTCAGAGGTTGACAGAAAAGCAGAGCGCCCTCAGTCCAGCC 1020
 Db 1549 GAACCTGTCCCTGCTGTCTCAGAGGTTGACAGAAAAGCAGAGCGCCCTCAGTCCAGCC 1608
 QY 1021 TGTGCGAATCTGTACTCTCAGAGGCAAGCAAAAGGCGCTGATCCCGCAGCGTG 1080
 Db 1609 TGTGCGAATCTGTACTCTCAGAGGCAAGCAAAAGGCGCTGATCCCGCAGCGTG 1668
 QY 1081 CCGAGCGTCCCGCAGCGTGTGCTGTGAAGAGCAGCGCTGTGTGAAGAGCGCTG 1140
 Db 1669 CCGAGCGTCCCGCAGCGTGTGCTGTGAAGAGCAGCGCTGTGTGAAGAGCGCTG 1728
 QY 1141 AGTGGGCTGACCTGTCTGCGAGCAGGCTGGAAGAGCAATTAAGTCAAGCTTCTCTTC 1200
 Db 1200 AGTGGGCTGACCTGTCTGCGAGCAGGCTGGAAGAGCAATTAAGTCAAGCTTCTCTTC

Db 1729 AGTGGGCTGACCTGTCTGCGAGCAGGCTGGAAGAGCAATTAAGTCAAGCTTCTTC 1788
 QY 1201 TCTCTGATATCAATCAGTTTCAATATCAGCCAGCATGTGAGATCTTACATGCTTC 1260
 Db 1789 TCTCTGATATCAATCAGTTTCAATATCAGCCAGCATGTGAGATCTTACATGCTTC 1848
 QY 1261 TCTCTGATATCAATCAGTTTCAATATCAGCCAGCATGTGAGATCTTACATGCTTC 1320
 Db 1849 TCTCTGATATCAATCAGTTTCAATATCAGCCAGCATGTGAGATCTTACATGCTTC 1908
 QY 1321 AAGCTATCCAGTTCTCCCTGTTTCAAGAACTATGGAAGAGTCTCCGAAACAGTCT 1380
 Db 1909 AAGCTATCCAGTTCTCCCTGTTTCAAGAACTATGGAAGAGTCTCCGAAACAGTCT 1968
 QY 1381 GATTAAGAGAGAGCAGCATTCCTCCAGAACTTCAAGAACTTCAAGAACTTCAAGAACT 1440
 Db 1969 GATTAAGAGAGAGCAGCATTCCTCCAGAACTTCAAGAACTTCAAGAACTTCAAGAACT 2028
 QY 1441 AGCAAGGATTTGATTCGATCAGAACGAGAGAGTGGCAAGGCTTCAAGAGTCTTCTT 1500
 Db 2029 AGCAAGGATTTGATTCGATCAGAACGAGAGAGTGGCAAGGCTTCAAGAGTCTTCTT 2088
 QY 1501 TCTTCAATGATTTGAAGTGGAGAGCTGTGAAGCAATTAACAACAGGCTTCTTCTTCTT 1560
 Db 2089 TCTTCAATGATTTGAAGTGGAGAGCTGTGAAGCAATTAACAACAGGCTTCTTCTTCTT 2148
 QY 1561 CTTTCCAGCAGCAGCAGCACTCAAGAGTGTGCTGCGCTTGAAGGCTGCGAC 1620
 Db 2149 CTTTCCAGCAGCAGCAGCACTCAAGAGTGTGCTGCGCTTGAAGGCTGCGAC 2208
 QY 1621 TCGGATATCTTGGCCCCCAGACTCTTACCCCTTCCCTGACCAAGCTGATATTTTGGC 1680
 Db 2209 TCGGATATCTTGGCCCCCAGACTCTTACCCCTTCCCTGACCAAGCTGATATTTTGGC 2268
 QY 1681 ACGAGTCTTCACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 Db 2269 ACGAGTCTTCACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2328
 QY 1741 GCTTCAAGTCTGAGCAGCTTCCCACTTGGGAGCAACCAAGTCTTCTGCTGCTGCTGCT 2388
 Db 2329 GCTTCAAGTCTGAGCAGCTTCCCACTTGGGAGCAACCAAGTCTTCTGCTGCTGCTGCT 2448
 QY 1801 CAGAAAGCAGTGAAGAGCTGATCTGCGGAGAGCTGAGTGAAGAGAGCCCTTTGAA 1860
 Db 2389 CAGAAAGCAGTGAAGAGCTGATCTGCGGAGAGCTGAGTGAAGAGAGCCCTTTGAA 2448
 QY 1861 AAGCAGTTTAAAGCAGAACTGCAATGGAATTTTGAAGAGATCATGTCAAGAAC 1920
 Db 2449 AAGCAGTTTAAAGCAGAACTGCAATGGAATTTTGAAGAGATCATGTCAAGAAC 2508
 QY 1921 AGGTCAAGGAGAGAGTGGGAAAGTGGGAGTCAAGTCTTCTTCTGCGGAGAGTGA 1980
 Db 2509 AGGTCAAGGAGAGAGTGGGAAAGTGGGAGTCAAGTCTTCTTCTGCGGAGAGTGA 2568
 QY 1981 ATCATTTGAAGTCTCC 1995
 Db 2569 ATCATTTGAAGTCTCC 2583

RESULT 15
 ABV21316 standard; cDNA; 5145 BP.
 ID ABV21316.
 AC ABV21316.
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 21307.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.


```
Db      2089  TCTCCAGTGCATCGAATGGAGAGTGGAGACAAATTACCAACAGCTTCTTTTCGGC 2148
Oy      1561  CTTTCCACCAAGCCGACGACACTTCACGAAGTCTGTGGCCTTGAGCCCTTAAGGGCTGGGAC 1620
Db      2149  CTTTCCACCAAGCCGACGACACTTCACGAAGTCTGTGGCCTTGAGCCCTTAAGGGCTGGGAC 2208
Oy      1621  TCGGATATCTTGGGCCCCGAGACCTTACCCCTTCCCTGACACGACGCTGTATTGTC 1680
Db      2209  TCGGATATCTTGGGCCCCGAGACCTTACCCCTTCCCTGACACGACGCTGTATTGTC 2268
Oy      1681  ACAGAGTCTTCACACTTCTACTCTGCTCAGCCATCTACGAGGCACTGCGAGTTACTCT 1740
Db      2269  ACAGAGTCTTCACACTTCTACTCTGCTCAGCCATCTACGAGGCACTGCGAGTTACTCT 2328
Oy      1741  GCCTACAGCTGCAGCCAGCTGCCCACTTGGCCGAGACCAAGTCTATTCTGTGCGAGCGG 1800
Db      2329  GCCTACAGCTGCAGCCAGCTGCCCACTTGGCCGAGACCAAGTCTATTCTGTGCGAGCGG 2388
Oy      1801  CAGAACCCAAAGTACAGAGCTGACTCGCGCGGAGCTGGCATGAAAGAGAGCCCTTGAA 1860
Db      2389  CAGAACCCAAAGTACAGAGCTGACTCGCGCGGAGCTGGCATGAAAGAGAGCCCTTGAA 2448
Oy      1861  AAGCAGTTTAAAGCGAGAAGCTGCCAATTGGAATTGGAAGAGCATCATGTCAGAGAAC 1920
Db      2449  AAGCAGTTTAAAGCGAGAAGCTGCCAATTGGAATTGGAAGAGCATCATGTCAGAGAAC 2508
Oy      1921  AGGTACCGGGAAGAGCTGGGGAAGAGTGGGAGTCACTAGCTTTTGGGCAAGCATGAA 1980
Db      2509  AGGTACCGGGAAGAGCTGGGGAAGAGTGGGAGTCACTAGCTTTTGGGCAAGCATGAA 2568
Oy      1981  ATCATTGAGGTCTCC 1995
Db      2569  ATCATTGAGGTCTCC 2583
```

Search completed: June 21, 2004, 18:35:44
Job time : 1167.45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 18:08:06 ; Search time 1191.09 Seconds
(without alignments)
7673.060 Million cell updates/sec

Title:	US-10-029-345A-108_COPY_538_2532
Perfect score:	1995
Sequence:	1 atggccatcgatgatctg.....tggaaatcatctgagctctcc 1995

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
Searched:      3017426 segs, 2290544650 residues
Total number of hits satisfying chosen parameters: 6034852
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA: *

- 1: /cgn2_6/prodataa1/1/pubnba/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodataa1/1/pubnba/PCOT_NEW_PUB.seq.*
- 3: /cgn2_6/prodataa1/1/pubnba/US06_NEW_PUB.seq.*
- 4: /cgn2_6/prodataa1/1/pubnba/US06_PUBCOMB.seq.*
- 5: /cgn2_6/prodataa1/1/pubnba/US07_NEW_PUB.seq.*
- 6: /cgn2_6/prodataa1/1/pubnba/PCOTUS_PUBCOMB.seq.*
- 7: /cgn2_6/prodataa1/1/pubnba/US08_NEW_PUB.seq.*
- 8: /cgn2_6/prodataa1/1/pubnba/US08_PUBCOMB.seq.*
- 9: /cgn2_6/prodataa1/1/pubnba/US09_PUBCOMB.seq.*
- 10: /cgn2_6/prodataa1/1/pubnba/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/prodataa1/1/pubnba/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/prodataa1/1/pubnba/US09_NEW_PUB.seq.*
- 13: /cgn2_6/prodataa1/1/pubnba/US10_NEW_PUB.seq.*
- 14: /cgn2_6/prodataa1/1/pubnba/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/prodataa1/1/pubnba/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/prodataa1/1/pubnba/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/prodataa1/1/pubnba/US10_NEW_PUB.seq.*
- 18: /cgn2_6/prodataa1/1/pubnba/US06_NEW_PUB.seq.*
- 19: /cgn2_6/prodataa1/1/pubnba/US06_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1991.8	99.8	1998	9	US-09-816-494-3	Sequence 3, Appl1
2	1991.8	99.8	1998	16	US-10-377-072-27	Sequence 27, Appl1
3	1991.8	99.8	3059	17	US-10-257-026-1	Sequence 1, Appl1
4	1991.8	99.8	3466	9	US-09-964-277-1	Sequence 1, Appl1
5	1991.8	99.8	3544	9	US-09-816-494-1	Sequence 1, Appl1
6	1991.8	99.8	3544	16	US-10-377-072-25	Sequence 25, Appl1
7	1991.8	99.8	3625	13	US-10-425-114-26234	Sequence 26234, Appl1
8	1991.8	99.8	3766	17	US-10-343-357-17	Sequence 17, Appl1
9	1991.8	99.8	4790	13	US-10-164-553-115	Sequence 115, Appl1
10	1990.2	99.8	2732	13	US-10-168-506-2	Sequence 2, Appl1
11	1988.6	99.7	2102	16	US-10-094-749-673	Sequence 673, Appl1
12	1988.6	99.7	2966	13	US-10-296-115-520	Sequence 520, Appl1
13	1982.2	98.4	2071	13	US-10-072-012-257	Sequence 257, Appl1
14	1906.8	95.6	2200	13	US-10-072-012-235	Sequence 235, Appl1

15	1653.8	82.9	3332	9	US-09-964-277-20	Sequence 20, Appl
16	879.4	44.1	1916	16	US-10-108-260A-2429	Sequence 2429, Ap
17	382	19.1	2453	14	US-10-005-858-1	Sequence 1, Appl
18	357.8	14.7	2476	13	US-10-220-170-28	Sequence 28, Appl
19	294.2	14.7	2039	13	US-10-072-012-265	Sequence 265, App
20	272.4	13.7	877	9	US-09-764-851-158	Sequence 158, App
21	223	11.2	277	17	US-10-637-855-29	Sequence 29, Appl
22	223	11.2	279	16	US-10-305-720-91	Sequence 91, Appl
23	217	10.9	246	17	US-10-257-026-3	Sequence 3, Appl
24	180.6	9.1	424	9	US-09-783-590-8120	Sequence 8120, Ap
25	124.6	6.2	1337	9	US-09-964-899-46	Sequence 46, Appl
26	124.6	6.2	1449	12	US-10-052-482-192	Sequence 192, App
27	124.6	6.2	1635	17	US-10-648-593-27	Sequence 27, Appl
28	124.6	6.2	1830	15	US-10-346-356-1	Sequence 1, Appl
29	124.6	6.2	1909	12	US-10-052-482-191	Sequence 191, App
30	124.6	6.2	1909	13	US-10-058-270A-129	Sequence 129, App
31	124.6	6.2	3286	10	US-09-971-352-9	Sequence 9, Appl
32	115	5.8	1452	12	US-10-052-482-189	Sequence 189, App
33	115	5.8	1977	12	US-10-052-482-188	Sequence 188, App
34	110.6	5.5	1356	15	US-10-184-832-6	Sequence 6, Appl
35	110.6	5.5	2741	15	US-10-184-832-4	Sequence 4, Appl
36	106.8	5.4	2303	9	US-09-880-107-3849	Sequence 3849, Ap
37	106.8	5.4	2303	10	US-09-892-146-3	Sequence 3, Appl
38	106.8	5.4	2303	13	US-10-342-867-603	Sequence 603, App
39	106.8	5.4	2303	13	US-10-172-118-603	Sequence 603, App
40	106.8	5.4	2338	13	US-10-363-616-7	Sequence 7, Appl
41	106.6	5.3	1143	13	US-10-184-832-3	Sequence 3, Appl
42	106.6	5.3	1574	13	US-10-296-115-440	Sequence 440, App
43	106.6	5.3	2650	15	US-10-184-832-1	Sequence 1, Appl
44	103.4	5.2	2104	12	US-10-152-319A-2041	Sequence 2041, App
45	103.4	5.2	2104	16	US-10-388-934-654	Sequence 654, App

ALIGNMENTS

```

RESULT 1
US-09-816-494-3
: Sequence 3, Application US/09816494
: Patent No. US20020034807A1
: GENERAL INFORMATION:
: APPLICANT: Meyers, Rachel A.
: TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
: TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
: FILE REFERENCE: 10448-030002
: CURRENT APPLICATION NUMBER: US/09/816,494
: PRIOR FILING DATE: 2001-03-23
: PRIOR APPLICATION NUMBER: US 60/191,858
: PRIOR FILING DATE: 2000-03-24
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 1998
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-816-494-3

```

Query Match	99.8%	Score 1991.8;	DB 9;	Length 1998;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1993; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

Qy	Db	Qy	Db	Qy	Db	Qy	Db
1	1	61	61	121	121	121	121
ATGGCCCATGATGATTTGGAACTCAAAATTGTTACTGAGAGGTTGGTGGCTCTGCTGGAA	ATGCCCCATGAGATGATTTGGAACTCAAAATTGTTACTGAGAGGTTGGTGGCTCTGCTGGAA	AGTGAACGGAAAAAGTCTCTTAATTGATAGCCGGCAATTTGTGGAATAACAATATATCC	AGTGAACGGAAAAAGTCTCTTAATTGATAGCCGGCAATTTGTGGAATAACAATATATCC	CACATTTTGGAAAGCCATTAATATATCAATGCTCTCAAGCTTTATGAACGAAGTTTGCACACAG	CACATTTTGGAAAGCCATTAATATATCAATGCTCTCAAGCTTTATGAACGAAGTTTGCACACAG	CACATTTTGGAAAGCCATTAATATATCAATGCTCTCAAGCTTTATGAACGAAGTTTGCACACAG	CACATTTTGGAAAGCCATTAATATATCAATGCTCTCAAGCTTTATGAACGAAGTTTGCACACAG
60	60	120	120	180	180	180	180

QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGATTCAAGCAAACTTAAGTTGACATTGAT 240
DB |||||
DB 181 GACAAAGTGTATTAACAGAGCTCATCCAGATTCAAGCAAACTTAAGTTGACATTGAT 240
QY 241 TGCAGTCAAGAGGTTGATGTTTAAAGATCAAGGCTCCGAAAGTGTGACATTGAT 240
DB |||||
DB 241 TGCAGTCAAGAGGTTGATGTTTAAAGATCAAGGCTCCGAAAGTGTGACATTGAT 240
QY 301 GACTGTTTCTCACTGTACTTCTGGGTTAACTGAGAAAGCTTCAACTCTGTTCACCTG 360
DB |||||
DB 301 GACTGTTTCTCACTGTACTTCTGGGTTAACTGAGAAAGCTTCAACTCTGTTCACCTG 360
QY 361 CTTCAGAGTGGGTTTGTCTGAGTTCTCTGTTTCTGCTGGCTCTGAGAAAGAAATCC 420
DB |||||
DB 361 CTTCAGAGTGGGTTTGTCTGAGTTCTCTGTTTCTGCTGGCTCTGAGAAAGAAATCC 420
QY 421 ACTCTAGTCCCTACCTGACATTCTCAGCCCTTCTTACCTGTTGCAACATTTGGGCAAC 480
DB |||||
DB 421 ACTCTAGTCCCTACCTGACATTCTCAGCCCTTCTTACCTGTTGCAACATTTGGGCAAC 480
QY 481 GGAATTTCTTCCCAATCTTTATCTTGCTGCTGCGAGAGATGTCTCAACAGAGACTGATA 540
DB |||||
DB 481 GGAATTTCTTCCCAATCTTTATCTTGCTGCTGCGAGAGATGTCTCAACAGAGACTGATA 540
QY 541 CAGCAGATGGGATTTGTTATGTTTAAAGTCACTTACCTGTTCAAAAGCTTGAATT 600
DB |||||
DB 541 CAGCAGATGGGATTTGTTATGTTTAAAGTCACTTACCTGTTCAAAAGCTTGAATT 600
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGATGACAGCTTTTGTGAAATTTTG 660
DB |||||
DB 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGATGACAGCTTTTGTGAAATTTTG 660
QY 661 CCGTGGTTGCAAAATCAGTATTTTCAATGAGAAACCAAAAGCTTCCAAATGATGTT 720
DB |||||
DB 661 CCGTGGTTGCAAAATCAGTATTTTCAATGAGAAACCAAAAGCTTCCAAATGATGTT 720
QY 721 CTAGTGCATCTTTTAACTGGGATCTCCGCTCCGCAACATCTGCTCAACATCTG 780
DB |||||
DB 721 CTAGTGCATCTTTTAACTGGGATCTCCGCTCCGCAACATCTGCTCAACATCTG 780
QY 781 AAGAGATGACATGTCTTTAGATGAAAGCTTAAGTTTGTAAAGAAAGAAAGCTTACT 840
DB |||||
DB 781 AAGAGATGACATGTCTTTAGATGAAAGCTTAAGTTTGTAAAGAAAGAAAGCTTACT 840
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGAAAGCTTACT 900
DB |||||
DB 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGAAAGCTTACT 900
QY 901 CAGACTGAGCATCAGGGCCAAAGAGCAACTCAAGCTGTGACCTTGAGAGAGCCAAAT 960
DB |||||
DB 901 CAGACTGAGCATCAGGGCCAAAGAGCAACTCAAGCTGTGACCTTGAGAGAGCCAAAT 960
QY 961 GAACTGTCTCTGCTGTCTCAGAGGGTGGACAGAAAGAGAGAGCCCTCAAGTCCACC 1020
DB |||||
DB 961 GAACTGTCTCTGCTGTCTCAGAGGGTGGACAGAAAGAGAGAGCCCTCAAGTCCACC 1020
QY 1021 TGTGCGACTGTCTACTCAGAGGCAAGAGCAAAAGCCCTGTGATCCCGCAGCGTG 1080
DB |||||
DB 1021 TGTGCGACTGTCTACTCAGAGGCAAGAGCAAAAGCCCTGTGATCCCGCAGCGTG 1080
QY 1081 CCCAGCGGCCAGGCTGCAAGCTGTCTGTTAGAGAGAGCCCGCTGGTACAGGGGCTC 1140
DB |||||
DB 1081 CCCAGCGGCCAGGCTGCAAGCTGTCTGTTAGAGAGAGCCCGCTGGTACAGGGGCTC 1140
QY 1141 AGTGGGCTGCACTGTCCGAGAGAGGCTGGAAGAGAGCAATAGCTCAAGGCTTCTTC 1200
DB |||||
DB 1141 AGTGGGCTGCACTGTCCGAGAGAGGCTGGAAGAGAGCAATAGCTCAAGGCTTCTTC 1200
QY 1201 TCTGTGATATCAATCAGTTTATATTCAGCCAGATGGAGCATCTTCAATGAGCTTC 1260
DB |||||
DB 1201 TCTGTGATATCAATCAGTTTATATTCAGCCAGATGGAGCATCTTCAATGAGCTTC 1260
QY 1261 TCTCATCAGAGATGCTTTGAAATCTTCACTTCTGATGGAGCAAC 1320
DB |||||
DB 1261 TCTCATCAGAGATGCTTTGAAATCTTCACTTCTGATGGAGCAAC 1320
QY 1321 AAGCTATGCAAGTCTCCCTCTGTGAGAACTTATGAGAGCACTCCGAAACCAAGCTCT 1380
DB |||||
DB 1321 AAGCTATGCAAGTCTCCCTCTGTGAGAACTTATGAGAGCACTCCGAAACCAAGCTCT 1380
QY 1381 GATTAAGAGAGAGAGAGAGATCCCAAGAGAGCTGAGAGCCGAGGCTTTAGACAGCAG 1440
DB |||||
DB 1381 GATTAAGAGAGAGAGAGAGATCCCAAGAGAGCTGAGAGCCGAGGCTTTAGACAGCAG 1440
QY 1441 AGCAAGCATTTGCAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB |||||
DB 1441 AGCAAGCATTTGCAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 TCTCCACTGCAATCGAATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB |||||
DB 1501 TCTCCACTGCAATCGAATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1561 CTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB |||||
DB 1561 CTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 TCGGATATCTTGGGCCCCAGAGCTTCACTTCCCTTCCCTTCCCTTCCCTTCCCT 1680
DB |||||
DB 1621 TCGGATATCTTGGGCCCCAGAGCTTCACTTCCCTTCCCTTCCCTTCCCTTCCCT 1680
QY 1681 ACAGAGTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1740
DB |||||
DB 1681 ACAGAGTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1740
QY 1741 GCTTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
DB |||||
DB 1741 GCTTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1801 CAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
DB |||||
DB 1801 CAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
QY 1861 AAGCAGTTTAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
DB |||||
DB 1861 AAGCAGTTTAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 1921 AGGTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
DB |||||
DB 1921 AGGTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
QY 1981 ATCATTTAGAGTCTTC 1995
DB |||||
DB 1981 ATCATTTAGAGTCTTC 1995
RESULT 2
US-10-377-072-27
Sequence 27, Application US/10377072
Publication No. US20040009501A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals Inc.
APPLICANT: Curtis, Rory A.J.
APPLICANT: Logan, Thomas Joseph
APPLICANT: Glucksmann, Maria A.
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Mlyoung
APPLICANT: Teal, Fong-yiing
TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
FILE REFERENCE: MP103-0180WIM
CURRENT APPLICATION NUMBER: US/10/377, 072
PRIOR FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: US 09/895, 860

;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: US 60/215,370
;; PRIOR FILING DATE: 2000-06-29
;; PRIOR APPLICATION NUMBER: US 09/723,806
;; PRIOR FILING DATE: 2000-11-28
;; PRIOR APPLICATION NUMBER: US 60/187,455
;; PRIOR FILING DATE: 2000-03-07
;; PRIOR APPLICATION NUMBER: US 09/843,297
;; PRIOR FILING DATE: 2001-04-25
;; PRIOR APPLICATION NUMBER: US 60/199,801
;; PRIOR FILING DATE: 2000-04-26
;; PRIOR APPLICATION NUMBER: US 09/861,801
;; PRIOR FILING DATE: 2001-05-21
;; PRIOR APPLICATION NUMBER: US 60/205,508
;; PRIOR FILING DATE: 2000-05-19
;; PRIOR APPLICATION NUMBER: US 09/816,494
;; PRIOR FILING DATE: 2001-03-23
;; PRIOR APPLICATION NUMBER: US 09/815,419
;; Remaining Prior Application data removed - See File Wrapper or PAM.
;; NUMBER OF SEQ ID NOS: 114
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 27
;; LENGTH: 1998
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)...(1998)
US-10-377-072-27

Query Match 99.8%; Score 1991.8; DB 16; Length 1998;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCATGAGATGATTTGGAACCAATTGTTACTGAGAGAGTTGGCTCTGCTGGAA 60
Db 1 ATGGCCATGAGATGATTTGGAACCAATTGTTACTGAGAGAGTTGGCTCTGCTGGAA 60
Qy 61 AGTGAACGGAAGAGTGTCTGCTAATTTGATGATGCGGCAATTTGGAAATACATCC 120
Db 61 AGTGAACGGAAGAGTGTCTGCTAATTTGATGATGCGGCAATTTGGAAATACATCC 120
Qy 121 CACATTTGGAAGCCATTAATATCACTGCTCCAAAGTTATGAAGCAAGTTGCAACAG 180
Db 121 CACATTTGGAAGCCATTAATATCACTGCTCCAAAGTTATGAAGCAAGTTGCAACAG 180
Qy 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATGAT 240
Db 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATGAT 240
Qy 241 TGCAGTCAGAGGTTGATGTTTACGATCAAGCTCCCAAGATGTGCTCTCTTCA 300
Db 241 TGCAGTCAGAGGTTGATGTTTACGATCAAGCTCCCAAGATGTGCTCTCTTCA 300
Qy 301 GACTGTTTCTCACTGATCTTTGGGTAACCTGGAAGAGCTTCACTCTGTTCACTG 360
Db 301 GACTGTTTCTCACTGATCTTTGGGTAACCTGGAAGAGCTTCACTCTGTTCACTG 360
Qy 361 CTTCGAGGTGGGTTTGTGATGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGAAATCC 420
Db 361 CTTCGAGGTGGGTTTGTGATGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGAAATCC 420
Qy 421 ACTCTAGTCCCTACCTGACATTTCTGAGCTTGTCTTACCTGTTGCCAATTTGGCCAAC 480
Db 421 ACTCTAGTCCCTACCTGACATTTCTGAGCTTGTCTTACCTGTTGCCAATTTGGCCAAC 480
Qy 481 CGAATTTCTTCCCAATCTTTATCTTGGCTGCAAGGAATGTCCTCAACAAGAGCTGATA 540
Db 481 CGAATTTCTTCCCAATCTTTATCTTGGCTGCAAGGAATGTCCTCAACAAGAGCTGATA 540
Qy 541 CAGCAGATGGGATTTGTTATGTTAAATGCGAGTATACCTGTCCAAAGCTGACTTT 600
Db 541 CAGCAGATGGGATTTGTTATGTTAAATGCGAGTATACCTGTCCAAAGCTGACTTT 600

Db 541 CAGCAGATGGGATTTGTTATGTTAAATGCGAGCAATACCTGTCCAAAGCTGACTTT 600
Qy 601 ATCCCGAGTCTCATTTCCGAGGTGCTGGAATGACAGCTTTGTGGAATAATTTTG 660
Db 601 ATCCCGAGTCTCATTTCCGAGGTGCTGGAATGACAGCTTTGTGGAATAATTTTG 660
Qy 661 CCGTGTGGAACAAATCAGTATGTTTCAATTTGAGAAAGCAAAAGCTTCAATGATGTG 720
Db 661 CCGTGTGGAACAAATCAGTATGTTTCAATTTGAGAAAGCAAAAGCTTCAATGATGTG 720
Qy 721 CTAGTCACTGTTTACCTGAGATCTCCGCTCCGCCCAATGCTATCCGCTATCATATG 780
Db 721 CTAGTCACTGTTTACCTGAGATCTCCGCTCCGCCCAATGCTATCCGCTATCATATG 780
Qy 781 AAGAGATGACATGCTTTAGATGAAGCTTCAATTTGTGAAAGAAAGAAAGCACTACT 840
Db 781 AAGAGATGACATGCTTTAGATGAAGCTTCAATTTGTGAAAGAAAGAAAGCACTACT 840
Qy 841 ATATCTCAAACTTCAATTTTCTGGCCCAATCTCTGACTATGAGAGAAATTAAGAAC 900
Db 841 ATATCTCAAACTTCAATTTTCTGGCCCAATCTCTGACTATGAGAGAAATTAAGAAC 900
Qy 901 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTCTGCACTGAGAGCCAAAT 960
Db 901 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTCTGCACTGAGAGCCAAAT 960
Qy 961 GAACTGTCTCTGCTCTCTCAAGAGGTGACAGAAAGGAGAGCGCCCTAGTCCAGCC 1020
Db 961 GAACTGTCTCTGCTCTCTCAAGAGGTGACAGAAAGGAGAGCGCCCTAGTCCAGCC 1020
Qy 1021 TGTGCGGACTCTGCTACTCAGAGGAGAGCAAAAGCCGCTGATCCGCGCAGAGTG 1080
Db 1021 TGTGCGGACTCTGCTACTCAGAGGAGAGCAAAAGCCGCTGATCCGCGCAGAGTG 1080
Qy 1081 CCCAGGCTGCCAGCTGAGAGCGCTGCTGTTAGAGAGCAAGCCGCTGATCAAGGCTTC 1140
Db 1081 CCCAGGCTGCCAGCTGAGAGCGCTGCTGTTAGAGAGCAAGCCGCTGATCAAGGCTTC 1140
Qy 1141 AGTGGCTGACCTGCTCCGAGCAGGCTGATGAGAGAGCAAGCAATGCTCAAGCTTCTTC 1200
Db 1141 AGTGGCTGACCTGCTCCGAGCAGGCTGATGAGAGAGCAATGCTCAAGCTTCTTC 1200
Qy 1201 TCTCTGATATCAAAATCAGTTCATATTCAGCCAGATGAGCATCTTACATGCTTC 1260
Db 1201 TCTCTGATATCAAAATCAGTTCATATTCAGCCAGATGAGCATCTTACATGCTTC 1260
Qy 1261 TCTCTATCAGAGATGCTTTGAAATCAAACTTCACTGATGAGAGCAAC 1320
Db 1261 TCTCTATCAGAGATGCTTTGAAATCAAACTTCACTGATGAGAGCAAC 1320
Qy 1321 AAGCTATGCGAGTTCCTCTGTTAGAGAGCTATGGAAGAGCTCCGGAACCAAGTCT 1380
Db 1321 AAGCTATGCGAGTTCCTCTGTTAGAGAGCTATGGAAGAGCTCCGGAACCAAGTCT 1380
Qy 1381 GATTAAGAGAGAGCGAGATCCCAAGAGCTCAGACGCGCTTCAAGAGCAG 1440
Db 1381 GATTAAGAGAGAGCGAGATCCCAAGAGCTCAGACGCGCTTCAAGAGCAG 1440
Qy 1441 AGCAAGCGATGATTCGCTCAGAACCAAGAGTGGCAAGCGCCCAAGAGTCTCTTTTA 1500
Db 1441 AGCAAGCGATGATTCGCTCAGAACCAAGAGTGGCAAGCGCCCAAGAGTCTCTTTTA 1500
Qy 1501 TCTTCATCTGATGAGAGTGGAGCGTGGAGAAATTAACAACAGAGCTTCTTTCGAG 1560
Db 1501 TCTTCATCTGATGAGAGTGGAGCGTGGAGAAATTAACAACAGAGCTTCTTTCGAG 1560
Qy 1561 CTTTCAACAGCAGAGCACTCAGAAAGTCTGAGCTGAGCTTGAAGGCTGAGC 1620
Db 1561 CTTTCAACAGCAGAGCACTCAGAAAGTCTGAGCTGAGCTTGAAGGCTGAGC 1620
Qy 1621 TCGGATATCTGGCCCCCAGACCTTACCCCTTGAACAGAGCTGATATTTGGCC 1680
Db 1621 TCGGATATCTGGCCCCCAGACCTTACCCCTTGAACAGAGCTGATATTTGGCC 1680

Qy 1681 ACAGAGCTCTACACTTCTTACTCTGCTTCCAGCCATCTTACGAGGAGAGTCCAGTTACTCT 1740
Db 1681 ACAGAGCTCTACACTTCTTACTCTGCTTCCAGCCATCTTACGAGGAGAGTCCAGTTACTCT 1740
Qy 1741 GCTTACAGCTGAGCCAGCTGCGCACTTGGCGAGAACCAAGCTTCTTCTGTCGCGAGCGG 1800
Db 1741 GCTTACAGCTGAGCCAGCTGCGCACTTGGCGAGAACCAAGCTTCTTCTGTCGCGAGCGG 1800
Qy 1801 CAGAGCCAGTGAAGAGCTGACTCTGCGGCGAGCTGGCATGAGAGAGAGCCCTTTGAA 1860
Db 1801 CAGAGCCAGTGAAGAGCTGACTCTGCGGCGAGCTGGCATGAGAGAGAGCCCTTTGAA 1860
Qy 1861 AAGCATTAAACGAGAGCTGCAATGGAATTTGAGAGAGCATCTGTCAGAGAAC 1920
Db 1861 AAGCATTAAACGAGAGCTGCAATGGAATTTGAGAGAGCATCTGTCAGAGAAC 1920
Qy 1921 AGTCAAGGAAAGAGCTGGGAAAGTGGGCACTGCTTCTTCTGCGGAGCATGGAA 1980
Db 1921 AGTCAAGGAAAGAGCTGGGAAAGTGGGCACTGCTTCTTCTGCGGAGCATGGAA 1980
Qy 1981 ATCATTTAGGCTTCC 1995
Db 1981 ATCATTTAGGCTTCC 1995

RESULT 3

US-10-257-026-1
Sequence 1, Application US/10257026
Publication No. US20040086859A1
GENERAL INFORMATION:
APPLICANT: Merck Patent GmbH
TITLE OF INVENTION: New dual specificity phosphatase
FILE REFERENCE: DUSP10KDM5
CURRENT APPLICATION NUMBER: US/10/257.026
CURRENT FILING DATE: 2003-11-07
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 3059
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (127)..(2121)
US-10-257-026-1

Query Match 99.8%; Score 1991.8; DB 17; Length 3059;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATGGAACTCAAAATTGTTACTGAGAGTGGTGGCTTGGTGA 60
Db 127 ATGGCCCATGAGATGATGAACTCAAAATTGTTACTGAGAGTGGTGGCTTGGTGA 186
Qy 61 AGTGAACGGAAAAAGTGTCTGTAATTGATAGCCGSCCATTTGAGAAATCAATCATCC 120
Db 187 AGTGAACGGAAAAAGTGTCTGTAATTGATAGCCGSCCATTTGAGAAATCAATCATCC 246
Qy 121 CACATTTTGAAGCATTATATCAACGCTCCAAAGCTTATGAGAGGAAAGTGGCAACG 180
Db 247 CACATTTTGAAGCATTATATCAACGCTCCAAAGCTTATGAGAGGAAAGTGGCAACG 306
Qy 181 GACAAAGGTTAATTAACAGAGCTCATCAGCATTCAAGCGAAATATAGGTTACATTGAT 240
Db 307 GACAAAGGTTAATTAACAGAGCTCATCAGCATTCAAGCGAAATATAGGTTACATTGAT 366
Qy 241 TGCAGTCAAGAGTTGATTTTGAATCAAGTCCCAAGATGTTGCTTCTCTTCA 300
Db 367 TGCAGTCAAGAGTTGATTTTGAATCAAGTCCCAAGATGTTGCTTCTCTTCA 426
Qy 301 GACTGTTTCTCACTGTAATCTTGGGTAATCTGAGAGAGGCTTCAATCTGTTCACTG 360
Db 301 GACTGTTTCTCACTGTAATCTTGGGTAATCTGAGAGAGGCTTCAATCTGTTCACTG 360

Db 427 GACTGTTTCTCACTGTAATCTTGGGTAATCTGAGAGAGGCTTCAATCTGTTCACTG 486
Qy 361 CTTCAGAGTGGTTTTCAGATTCCTGTTGTTCCCTGGCCCTTGAAGAAATCC 420
Db 487 CTTCAGAGTGGTTTTCAGATTCCTGTTGTTCCCTGGCCCTTGAAGAAATCC 546
Qy 421 ACTCTAGTCCCTACCTGCAATTTCTCAGGCTTGTACTCTGTCGCAATTTGGGCCAAC 480
Db 547 ACTCTAGTCCCTACCTGCAATTTCTCAGGCTTGTACTCTGTCGCAATTTGGGCCAAC 606
Qy 481 CGAATTTCTCCCAATCTTTATCTTGGTCCAGGAGATGCTCCCAAGAGAGGCTGAT 540
Db 607 CGAATTTCTCCCAATCTTTATCTTGGTCCAGGAGATGCTCCCAAGAGAGGCTGAT 666
Qy 541 CAGCAAAATGGGAATGGTTATGTTAAATCCAGCTTACCTGTCCAAAGCTGACTTT 600
Db 667 CAGCAAAATGGGAATGGTTATGTTAAATCCAGCTTACCTGTCCAAAGCTGACTTT 726
Qy 601 ATCCCGAGTCAATTTCTGCGTGTGCTGTGATGACAGCTTTTGTGAAGAAATTTTG 660
Db 727 ATCCCGAGTCAATTTCTGCGTGTGCTGTGATGACAGCTTTTGTGAAGAAATTTTG 786
Qy 661 CCGTGTGGAACAATCAAGATTTCAATTGAGAAAGCAAAAGCTCCATGATGATGTT 720
Db 787 CCGTGTGGAACAATCAAGATTTCAATTGAGAAAGCAAAAGCTCCATGATGATGTT 846
Qy 721 CTAGTGACCTGTTTACTGGGATCTCCCGCTCCGCACTATGCTATGCTGATCATGAT 780
Db 847 CTAGTGACCTGTTTACTGGGATCTCCCGCTCCGCACTATGCTATGCTGATCATGAT 906
Qy 781 AAGAGATGAGACATGCTTTAGATGAGCTTACAGATTTTGTGAAGAAAGAAAGCTTACT 840
Db 907 AAGAGATGAGACATGCTTTAGATGAGCTTACAGATTTTGTGAAGAAAGAAAGCTTACT 966
Qy 841 ATATCTCCCAATCTTAAATTTCTGGGCCAACTCTGTGACTATGAGAAATTTAAGAC 900
Db 967 ATATCTCCCAATCTTAAATTTCTGGGCCAACTCTGTGACTATGAGAAATTTAAGAC 1026
Qy 901 CAGACTGAGACATGAGGCCCAAGAGCAAACTCAAGCTGTGCACTGAGAAAGCCAAAT 960
Db 1027 CAGACTGAGACATGAGGCCCAAGAGCAAACTCAAGCTGTGCACTGAGAAAGCCAAAT 1086
Qy 961 GAACCTGTCCCTGCTGTCTGAGAGGCTGAGACAGAAAGGAGAGCCCTCAGTCCACC 1020
Db 1087 GAACCTGTCCCTGCTGTCTGAGAGGCTGAGACAGAAAGGAGAGCCCTCAGTCCACC 1146
Qy 1021 TGTCCCACTCTGCTACCTCAGAGGCGAGCAAGAGCCCGGTGATCCGCGAGGCTG 1080
Db 1147 TGTCCCACTCTGCTACCTCAGAGGCGAGCAAGAGCCCGGTGATCCGCGAGGCTG 1206
Qy 1081 CCGAGCGTGCAGCGTGAAGCGCTGCTGTTAAGAGACAGCCCGCTGTGACAGGCGCTG 1140
Db 1207 CCGAGCGTGCAGCGTGAAGCGCTGCTGTTAAGAGACAGCCCGCTGTGACAGGCGCTG 1266
Qy 1141 AGTGGCTGCACTGTGCGCAGACAGGCTGGAAGACAAATAGCTCAAGCTTCTCTT 1200
Db 1267 AGTGGCTGCACTGTGCGCAGACAGGCTGGAAGACAAATAGCTCAAGCTTCTCTT 1326
Qy 1201 TCTCTGATATCAATCAATGTTTCAATTTCAAGCCGACATGCGACATCTTCAATGCTTC 1260
Db 1327 TCTCTGATATCAATCAATGTTTCAATTTCAAGCCGACATGCGACATCTTCAATGCTTC 1386
Qy 1261 TCTCTGATATCAATGCTTTGGAATTAATCAAACTTTCATCTCTGATGAGGACCAAC 1320
Db 1387 TCTCTGATATCAATGCTTTGGAATTAATCAAACTTTCATCTCTGATGAGGACCAAC 1446
Qy 1321 AAGCTATGCGAGTCTCCCTGTTGAGAAATCTGAGAGACATCTCCGAAACAGTCTT 1380
Db 1447 AAGCTATGCGAGTCTCCCTGTTGAGAAATCTGAGAGACATCTCCGAAACAGTCTT 1506
Qy 1381 GATTAAGAGAAAGCAGCATCCCAAGAAAGCTGAGACCGCAGGCTTGAAGACAGCAG 1440
Db 1507 GATTAAGAGAAAGCAGCATCCCAAGAAAGCTGAGACCGCAGGCTTGAAGACAGCAG 1566

QY 1441 AGAAGGATTTGATTTGGTTCAGAACGAGAGAGTGGCAACCGCCAGAGTCCCTTTTA 1500
DB 1567 AGCAAGGATTTGATTTGGTTCAGAACGAGAGAGTGGCAACCGCCAGAGTCCCTTTTA 1626
QY 1501 TCTCCATGTCATGGAAGTGGAGCGTGGAGACAAATTACCAACAGACTTCTTTTGGG 1560
DB 1627 TCTCCATGTCATGGAAGTGGAGCGTGGAGACAAATTACCAACAGACTTCTTTTGGG 1686
QY 1561 CTTTCCACGACGACGACGACCTTCAAGAGTCTGCTGGCTGGGCTTTAAGGCTGGAC 1620
DB 1687 CTTTCCACGACGACGACGACCTTCAAGAGTCTGCTGGCTGGGCTTTAAGGCTGGAC 1746
QY 1621 TCGGATATCTTGGCCCCCGACGACCTTCAACCTTCCCTGACGACGACGCTGTAATTTGGC 1680
DB 1747 TCGGATATCTTGGCCCCCGACGACCTTCAACCTTCCCTGACGACGACGCTGTAATTTGGC 1806
QY 1681 ACAGAGTCTTCAACCTTCTACTCTGCTCAGACCACTTCAAGAGGAGTGGCACTTCT 1740
DB 1807 ACAGAGTCTTCAACCTTCTACTCTGCTCAGACCACTTCAAGAGGAGTGGCACTTCT 1866
QY 1741 GCCTACAGTTCAGACGACGCTGCTCCACTTGGGAGACCAAGTCTATTTCTGCGCAGGCG 1800
DB 1867 GCCTACAGTTCAGACGACGCTGCTCCACTTGGGAGACCAAGTCTATTTCTGCGCAGGCG 1926
QY 1801 CAGAGGCAAGTTCAGAGAGTGAATCTGCGGCGGAGAGTGGATGAAGAGAGCCCTTTGAA 1860
DB 1927 CAGAGGCAAGTTCAGAGAGTGAATCTGCGGCGGAGAGTGGATGAAGAGAGCCCTTTGAA 1986
QY 1861 AACGAGTTTAAAGCAGAGAGCTGCGCAATGGAATTTGAGAGAGCATGTCAGAGAAC 1920
DB 1987 AACGAGTTTAAAGCAGAGAGCTGCGCAATGGAATTTGAGAGAGCATGTCAGAGAAC 2046
QY 1921 AGGTTCAGGGAAGAGTGGGGAAGTGGGAGTCAAGTCTTCTTGGGAGCATGGA 1980
DB 2047 AGGTTCAGGGAAGAGTGGGGAAGTGGGAGTCAAGTCTTCTTGGGAGCATGGA 2106
QY 1981 ATCATTGAGGCTCC 1995
DB 2107 ATCATTGAGGCTCC 2121

RESULT 4

US-09-964-277-1
; Sequence 1, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucite, Ralf M.
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-1

Query Match 99.84; Score 1991.8; DB 9; Length 3496;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGATGATTTGAACTCAATTTGTAATGAGAGTGGTGGCTGTGAGAA 60
DB 562 ATGGCCCATGATGATTTGAACTCAATTTGTAATGAGAGTGGTGGCTGTGAGAA 621
QY 61 AGTGAACGGAAGAGTGTGCTTAATTTGATAGCCGCGCATTTTGGAAATCAATACC 120
DB 622 AGTGAACGGAAGAGTGTGCTTAATTTGATAGCCGCGCATTTTGGAAATCAATACC 681

QY 121 CACATTTGGAGCCATTAATATCACTGCTCCAGACTTATGAGAGGAAGTTGCAACAG 180
DB 682 CACATTTGGAGCCATTAATATCACTGCTCCAGACTTATGAGAGGAAGTTGCAACAG 741
QY 181 GACAAAGTGTAAATTAAGAGCTCATTCAGCATTCAGCGAACAATAGTTGACATTTGAT 240
DB 742 GACAAAGTGTAAATTAAGAGCTCATTCAGCATTCAGCGAACAATAGTTGACATTTGAT 801
QY 241 TGCAGTCAGAGGTTGATTTTCAATCAAGTCTCCAGATTTGCTCTCTTCA 300
DB 802 TGCAGTCAGAGGTTGATTTTCAATCAAGTCTCCAGATTTGCTCTCTTCA 861
QY 301 GACTTTTCTCACTGTAATCTTGGGTAACCTGGAAGAGCTTCAACTCTGTCACCTG 360
DB 862 GACTTTTCTCACTGTAATCTTGGGTAACCTGGAAGAGCTTCAACTCTGTCACCTG 921
QY 361 CTTCAGAGTGGGTTTCTGAGTTCTCTGTTTCTGCTGCTCTGTAAGGAAATCC 420
DB 922 CTTCAGAGTGGGTTTCTGAGTTCTCTGTTTCTGCTGCTCTGTAAGGAAATCC 981
QY 421 ACTTATGCTCTTACCTGCAATTTCTAGCCTTGTACTCTGTCGAACATTTGGCCAAAC 480
DB 982 ACTTATGCTCTTACCTGCAATTTCTAGCCTTGTACTCTGTCGAACATTTGGCCAAAC 1041
QY 481 CGAATTTCTTCCCAATCTTATCTTGGCTGCAAGAGATGCTCTCAACAGAGCTGATA 540
DB 1042 CGAATTTCTTCCCAATCTTATCTTGGCTGCAAGAGATGCTCTCAACAGAGCTGATA 1101
QY 541 CAGCAGATGGAATGTTATGTTAAATGCGAGTATACCTGTCCAAAGCTGACTTT 600
DB 1102 CAGCAGATGGAATGTTATGTTAAATGCGAGTATACCTGTCCAAAGCTGACTTT 1161
QY 601 ATCCCGAGTCTCATTTCTCGCTGCTGCTGTAATGACAGCTTTTGTGAGAAATTTTG 660
DB 1162 ATCCCGAGTCTCATTTCTCGCTGCTGCTGTAATGACAGCTTTTGTGAGAAATTTTG 1221
QY 661 CCGTGTGGAACAATCAGATGTTTCAATGAGAAAGCAAAAGCTCAATGATGTT 720
DB 1222 CCGTGTGGAACAATCAGATGTTTCAATGAGAAAGCAAAAGCTCAATGATGTT 1281
QY 721 CTAGTCACTGTTTACCTGGGATCTCCGCTCGGCAACATGCTATGCTATCATCATG 780
DB 1282 CTAGTCACTGTTTACCTGGGATCTCCGCTCGGCAACATGCTATGCTATCATCATG 1341
QY 781 AAGAGATGACATGCTTTAATGAGAGCTTACAGATTTGTGAAAGAAAGAACTTACT 840
DB 1342 AAGAGATGACATGCTTTAATGAGAGCTTACAGATTTGTGAAAGAAAGAACTTACT 1401
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
DB 1402 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGATTAAGAAC 1461
QY 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTCTGCACTGAGAGAACCAAT 960
DB 1462 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTCTGCACTGAGAGAACCAAT 1521
QY 961 GAACTGCTGCTGCTGCTCTCAAGAGGTTGACAGAAAGGAGAGCGCCCTCACTCAACC 1020
DB 1522 GAACTGCTGCTGCTGCTCTCAAGAGGTTGACAGAAAGGAGAGCGCCCTCACTCAACC 1581
QY 1021 TGTGCGGACTCTGCTATCTCAAGAGGAGAGCAAAAGGCGGTGATCCCGCAGCGTG 1080
DB 1582 TGTGCGGACTCTGCTATCTCAAGAGGAGAGCAAAAGGCGGTGATCCCGCAGCGTG 1641
QY 1081 CCCAGGCTGCCAGCGCTGAGCGCTGCTGTTAAGAGACAGCCGCTGTGTAACAGCGCTC 1140
DB 1642 CCCAGGCTGCCAGCGCTGAGCGCTGCTGTTAAGAGACAGCCGCTGTGTAACAGCGCTC 1701
QY 1141 AGTGGCTGCACTGCTGCGCAACAGGCTGGAAGACAGCAATAGCTCAAGGCTCTTC 1200
DB 1702 AGTGGCTGCACTGCTGCGCAACAGGCTGGAAGACAGCAATAGCTCAAGGCTCTTC 1761
QY 1201 TCTCTGATATCAATCAGTTTCAATTTCAAGCAGCATGGCAGCATCTTACATGCTTC 1260

```
Db 1762 TCTGTGATATCAAAATCAATTTCAATATCAAGCAAGCAAGCATCTTAACATGGCTTC 1821
Qy 1261 TCTGTATGAGAAATGCTTTTGAATATCAAAACCTTCACTACTGTGATGGGAACAAC 1320
Db 1822 TCTGTATGAGAAATGCTTTTGAATATCAAAACCTTCACTACTGTGATGGGAACAAC 1881
Qy 1321 AAGCTATGCAATGCTTCCCTGTGTGAGAAATATCGAAGCAAGCTCCGAAACCAATCTCT 1380
Db 1882 AAGCTATGCAATGCTTCCCTGTGTGAGAAATATCGAAGCAAGCTCCGAAACCAATCTCT 1941
Qy 1381 GATTAGAGGAAGCAAGCATCTCCCAAGAACTGTGAGACCGCAGGCTTCAACACGACG 1440
Db 1942 GATTAGAGGAAGCAAGCATCTCCCAAGAACTGTGAGACCGCAGGCTTCAACACGACG 2001
Qy 1441 AGCAAGCATGTTGCAATGCTGTGAGAAACGAGAGCAAGCGCCCAAGAGTCCCTTTTA 1500
Db 2002 AGCAAGCATGTTGCAATGCTGTGAGAAACGAGAGCAAGCGCCCAAGAGTCCCTTTTA 2061
Qy 1501 TCTGTACTGATGGAAGTGGAGAGCTGTGAGAAACAATTAACAACGAGCTTCTTTTGGC 1560
Db 2062 TCTGTACTGATGGAAGTGGAGAGCTGTGAGAAACAATTAACAACGAGCTTCTTTTGGC 2121
Qy 1561 CTCTTCAACAGCAGCAGCACTCAAGAACTGTGAGCTGTGAGCTTAAAGAGCTGAC 1620
Db 2122 CTCTTCAACAGCAGCAGCACTCAAGAACTGTGAGCTGTGAGCTTAAAGAGCTGAC 2181
Qy 1621 TCGGATATCTGGGCCCCCGAGACCTGTACCCCTTCCCTGACCAAGCACTGTGATTTTGGC 1680
Db 2182 TCGGATATCTGGGCCCCCGAGACCTGTACCCCTTCCCTGACCAAGCACTGTGATTTTGGC 2241
Qy 1681 AAGAGTCTCTCAACCTTCTACTGTGCTGAGCAATCAAGAGCAAGCTGCAAGTACTCT 1740
Db 2242 AAGAGTCTCTCAACCTTCTACTGTGCTGAGCAATCAAGAGCAAGCTGCAAGTACTCT 2301
Qy 1741 GCTTACAGTGTGAGCAAGCTGCACTGTGAGAAACAATTAATCTGTGAGCAAGCTG 1800
Db 2302 GCTTACAGTGTGAGCAAGCTGCACTGTGAGAAACAATTAATCTGTGAGCAAGCTG 2361
Qy 1801 CAGAAAGCAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860
Db 2362 CAGAAAGCAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2421
Qy 1861 AAGCAAGTTTAAAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1920
Db 2422 AAGCAAGTTTAAAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2481
Qy 1921 AAGTCAAGGGAAGAGCTGGGGAAGTGGGCAAGTCAAGTCTTGGGCAAGTGAAG 1980
Db 2482 AAGTCAAGGGAAGAGCTGGGGAAGTGGGCAAGTCAAGTCTTGGGCAAGTGAAG 2541
Qy 1981 ATCATTTGAGGTCTCC 1995
Db 2542 ATCATTTGAGGTCTCC 2556
```

RESULT 5

```
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589)
US-09-816-494-1
```

```
Query Match 99.8%; Score 1991.8; DB 9; Length 3544;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 ATGGCCCATGAGATGATTTGAACTCAAAATTTGTAAGAGAGTGTGGTGTGCTGCTGAA 60
Db 589 ATGGCCCATGAGATGATTTGAACTCAAAATTTGTAAGAGAGTGTGGTGTGCTGCTGAA 648
Qy 61 AGTGGAAAGGAAAGAGTGTGCTAATTTGATGAGAGGAGGCTTTGTGAAATACATAATCC 120
Db 649 AGTGGAAAGGAAAGAGTGTGCTAATTTGATGAGAGGAGGCTTTGTGAAATACATAATCC 708
Qy 121 CAACATTTGGAAGCATTAATATCAATCTGCTCAAGCTTATGAGAGGAGTGTGCAAG 180
Db 709 CAACATTTGGAAGCATTAATATCAATCTGCTCAAGCTTATGAGAGGAGTGTGCAAG 768
Qy 181 GACAAAGTGTAAATTAAGAGCTATCCAGCAATTAAGGAAACATTAAGTGTGACATGAT 240
Db 769 GACAAAGTGTAAATTAAGAGCTATCCAGCAATTAAGGAAACATTAAGTGTGACATGAT 828
Qy 241 TGCAGTCAAGAGTGTGATTTTACATCAAAAGCTCCAAAGATGTGTGCTCTCTCTTCA 300
Db 829 TGCAGTCAAGAGTGTGATTTTACATCAAAAGCTCCAAAGATGTGTGCTCTCTCTTCA 888
Qy 301 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAAAGCTTCAACTGTGTAACCTG 360
Db 889 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAAAGCTTCAACTGTGTAACCTG 948
Qy 361 CTTCAGAGTGTGCTGAGTCTCTGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 949 CTTCAGAGTGTGCTGAGTCTCTGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008
Qy 421 ACTTGAAGTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 480
Db 1009 ACTTGAAGTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1068
Qy 481 GCAATTTCTCCCATCTTTATCTTGGTGTGCAAGAGTGTGCTGATCAAGAGCTGATCT 540
Db 1069 GCAATTTCTCCCATCTTTATCTTGGTGTGCAAGAGTGTGCTGATCAAGAGCTGATCT 1128
Qy 541 CAGCAAGATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 1129 CAGCAAGATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy 601 ATCCCGAGTCTGATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 1189 ATCCCGAGTCTGATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
Qy 661 CCGTGTGGAACAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 1249 CCGTGTGGAACAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308
Qy 721 CTAGTCACTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 1309 CTAGTCACTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1368
Qy 781 AAGAGATGAGATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 1369 AAGAGATGAGATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Qy 841 ATATCTCAAACTCAATTTTGTGGGCAACTCTGATGATGATGATGATGATGATGATGAT 1428
Db 1429 ATATCTCAAACTCAATTTTGTGGGCAACTCTGATGATGATGATGATGATGATGATGAT 1488
Qy 901 CAGACTGAGATCAGGCGCAAGCAAACTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
```

1489 CAGACTGAGATCAGGGCCAAAGACAACTCAAGCTGCTGACCTGAGAAACCAAT 1548
QY 961 GAACCTGCTCCTGCTGCTCAAGAGGTGACAGAAAAGCGAGAGCCCTTCACTGACCC 1020
DB 1549 GAACCTGCTCCTGCTGCTCAAGAGGTGACAGAAAAGCGAGAGCCCTTCACTGACCC 1608
QY 1021 TGTGCGGACTCTGCTACTCAAGAGGCGAGAGCAAAAGCCCGTGCATCCCGCAGAGGTG 1080
DB 1609 TGTGCGGACTCTGCTACTCAAGAGGCGAGAGCAAAAGCCCGTGCATCCCGCAGAGGTG 1668
QY 1081 CCGAGCGTGGCCAGCGCTGACAGCGCTGCTGCTGTAAGAGACAGCCGCTGTATACAGCCCTC 1140
DB 1669 CCGAGCGTGGCCAGCGCTGACAGCGCTGCTGCTGTAAGAGACAGCCGCTGTATACAGCCCTC 1728
QY 1141 AGTGGCTGCACTGCTGCTGCAAGAGGCTGGAAGACAGCAATTAAGCTTCAAGCTTCTTC 1200
DB 1729 AGTGGCTGCACTGCTGCTGCAAGAGGCTGGAAGACAGCAATTAAGCTTCAAGCTTCTTC 1788
QY 1201 TCTCTGATATCAATAGTTTCACTTCAAGCCAGATGCGAGATCTTACATGAGCTTC 1260
DB 1789 TCTCTGATATCAATAGTTTCACTTCAAGCCAGATGCGAGATCTTACATGAGCTTC 1848
QY 1261 TCTCATCAGAAAGATGCTTGAATATCACTCAACTTCACTGATGAGAGCCAC 1320
DB 1849 TCTCATCAGAAAGATGCTTGAATATCACTCAACTTCACTGATGAGAGCCAC 1908
QY 1321 AAGCTATGCAATGCTTCTCCTGTTAGAGAACTATCGAGCAACTCCGAAACAGTCT 1380
DB 1909 AAGCTATGCAATGCTTCTCCTGTTAGAGAACTATCGAGCAACTCCGAAACAGTCT 1968
QY 1381 GATPAGAGAGAGAGCAGATCCCAAGAGCTGAGACCGCCAGGCTTCAAGAGCAG 1440
DB 1969 GATPAGAGAGAGAGCAGATCCCAAGAGCTGAGACCGCCAGGCTTCAAGAGCAG 2028
QY 1441 AGCAAGGATGATGCTTGTGCTCAGAACAGAGAGAGTGGCAACCGCCAGAGTCCCTTTA 1500
DB 2029 AGCAAGGATGATGCTTGTGCTCAGAACAGAGAGAGTGGCAACCGCCAGAGTCCCTTTA 2088
QY 1501 TCTTCACTGATCGAAGTGGAGAGCTGAGAGCAATTAACAACACAGCTTCTTTTGGC 1560
DB 2089 TCTTCACTGATCGAAGTGGAGAGCTGAGAGCAATTAACAACACAGCTTCTTTTGGC 2148
QY 1561 CTTTCCACGAGCGAGCAAGCTTCAAGAGCTGCTGAGGCTTGAAGGCTGAGC 1620
DB 2149 CTTTCCACGAGCGAGCAAGCTTCAAGAGCTGCTGAGGCTTGAAGGCTGAGC 2208
QY 1621 TCGGATATCTTGGGCCCCCAGACCTTACCTTCCCTGACCAAGAGCTGATTTTGGC 1680
DB 2209 TCGGATATCTTGGGCCCCCAGACCTTACCTTCCCTGACCAAGAGCTGATTTTGGC 2268
QY 1681 ACAAGTCTTCACTTCTTCTGCTGCTGAGCTTCAAGCAATTAAGAGGAGTGCATCTCT 1740
DB 2269 ACAAGTCTTCACTTCTTCTGCTGCTGAGCTTCAAGCAATTAAGAGGAGTGCATCTCT 2328
QY 1741 GCGTACGCTGAGCGAGCGAGCTGAGCACTTGGGAGAACCAAGCTTATTTGGGAGGCG 1800
DB 2329 GCGTACGCTGAGCGAGCGAGCTGAGCACTTGGGAGAACCAAGCTTATTTGGGAGGCG 2388
QY 1801 CAGAAAGCAAGTGAAGAGTGAATCTCGCGGAGCTGAGCAAGAGAGAGAGCCCTTTGAA 1860
DB 2389 CAGAAAGCAAGTGAAGAGTGAATCTCGCGGAGCTGAGCAAGAGAGAGAGCCCTTTGAA 2448
QY 1861 AAGCAATTTAAACGAGAGCTGCAATGGAATTTGAGAGAGCAATGTCAGAGAAC 1920
DB 2449 AAGCAATTTAAACGAGAGCTGCAATGGAATTTGAGAGAGCAATGTCAGAGAAC 2508
QY 1921 AGGTCAAGGAGAGAGCTGAGGAGAGTGGGAGTGAAGTCTGTTGGGAGAGATGAA 1980
DB 2509 AGGTCAAGGAGAGAGCTGAGGAGAGTGGGAGTGAAGTCTGTTGGGAGAGATGAA 2568
QY 1981 ATCATTTAGAGTCTCC 1995
DB 2569 ATCATTTAGAGTCTCC 2583

RESULT 6
US-10-377-072-25
; Sequence 25, Application US/10377072
; Publication No. US2004009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Miyoung
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; FILE REFERENCE: M0103-0180NMIM
; CURRENT APPLICATION NUMBER: US/10/377, 072
; CURRENT FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; PRIOR FILING DATE: 2001-03-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PASCSEQ for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589) ... (2586)
US-10-377-072-25
Query Match 99.8%; Score 1991.8; DB 16; Length 3544;
Bseq Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATGGAATGGAATGTTACTGAGAGGTTGCTGCTGAGAA 60
DB 589 ATGGCCCATGAGATGATGGAATGGAATGTTACTGAGAGGTTGCTGCTGAGAA 648
QY 61 AGTGAACGGAAGAGTGTCTGCTAATTTGATAGCCGCGCATTTGTGAAATCAATACATCC 120
DB 649 AGTGAACGGAAGAGTGTCTGCTAATTTGATAGCCGCGCATTTGTGAAATCAATACATCC 708
QY 121 CACATTTGGAACCATTAATATCAATGCTCCAGGCTTATGAAGCGAAGTTGCAACAG 180
DB 709 CACATTTGGAACCATTAATATCAATGCTCCAGGCTTATGAAGCGAAGTTGCAACAG 768
QY 181 GACAAAGTTTAATTAAGAGCTCATTCAGCATTCAGCGAACAATAGTTGACATTTGAT 240
DB 769 GACAAAGTTTAATTAAGAGCTCATTCAGCATTCAGCGAACAATAGTTGACATTTGAT 828

QY 241 TCGAGTGAAGAGTTGTTAGTTTACGATCAAAAGCTCCAGAGTGTGCTCTCTTCA 300
Db 829 TCGAGTGAAGAGTTGTTAGTTTACGATCAAAAGCTCCAGAGTGTGCTCTCTTCA 888
QY 301 GACTGTTTTCTCACTGTACTCTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACCTG 360
Db 889 GACTGTTTTCTCACTGTACTCTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACCTG 948
QY 361 CTTCGAGTGGGTTTGTGATGTTCTCTGTTGTTTCCCTGACCTGTGAGAGAAATCC 420
Db 949 CTTCGAGTGGGTTTGTGATGTTCTCTGTTGTTTCCCTGACCTGTGAGAGAAATCC 1008
QY 421 ACTCTAGTCCCTACCTGATTTCTCAAGCTTGTCTTACCTGTGACCAATGGGCAACC 1068
Db 1009 ACTCTAGTCCCTACCTGATTTCTCAAGCTTGTCTTACCTGTGACCAATGGGCAACC 1068
QY 481 CGAATCTTCCCAATCTTAACTTGTGCTGCGACGAGATGCTCAACAGAGGCTGATA 540
Db 1069 CGAATCTTCCCAATCTTAACTTGTGCTGCGACGAGATGCTCAACAGAGGCTGATA 1128
QY 541 CAGCAGATGGGATGGTATATGTATTAATGACAGCTAACCTGTCCAAAGCTGACTTT 600
Db 1129 CAGCAGATGGGATGGTATATGTATTAATGACAGCTAACCTGTCCAAAGCTGACTTT 1188
QY 601 ATCCCGAGTCTCATTTCTCGGCTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 1248
Db 1189 ATCCCGAGTCTCATTTCTCGGCTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 1248
QY 661 CCGTGTGTGACCAATCAGTAGATTCTTGAAGAAAGCAAGCTCAATGATGTTT 720
Db 1249 CCGTGTGTGACCAATCAGTAGATTCTTGAAGAAAGCAAGCTCAATGATGTTT 720
QY 721 CTAGTGACCTGTTAGCTGGGATCTCCGCTCCGCCACATCTGCTATCGCTACATCATG 780
Db 1309 CTAGTGACCTGTTAGCTGGGATCTCCGCTCCGCCACATCTGCTATCGCTACATCATG 1368
QY 781 AAGAGATGGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 840
Db 1369 AAGAGATGGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 840
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGACCTTATGAGAGAAATTAAGAAC 1428
Db 1429 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGACCTTATGAGAGAAATTAAGAAC 1488
QY 901 CAGACTGAGATCAAGGAGCCAAAGACAACTCAAGCTGTGACCTGAGAGAGCAAT 960
Db 1489 CAGACTGAGATCAAGGAGCCAAAGACAACTCAAGCTGTGACCTGAGAGAGCAAT 1548
QY 961 GAACCTGTCTCTGCTCAAGGGGTGAGAGAAAGCCCTCAAGTCAACC 1020
Db 1549 GAACCTGTCTCTGCTCAAGGGGTGAGAGAAAGCCCTCAAGTCAACC 1608
QY 1021 TGTGCGACCTGTCTCAAGGACAGAGCAAAAGCCCTGTGACCTGAGAGAGCTG 1080
Db 1609 TGTGCGACCTGTCTCAAGGACAGAGCAAAAGCCCTGTGACCTGAGAGAGCTG 1668
QY 1081 CCCAGGTTCCAGGCTGAGCCGTGTGTTAGAGACAGCCCGCTGTGACAGGCTG 1140
Db 1669 CCCAGGTTCCAGGCTGAGCCGTGTGTTAGAGACAGCCCGCTGTGACAGGCTG 1728
QY 1141 AGTGGGCTGACCTGTCCGAGACAGGCTGAGAGCAAAAGCTCAAGGCTTCTT 1200
Db 1729 AGTGGGCTGACCTGTCCGAGACAGGCTGAGAGCAAAAGCTCAAGGCTTCTT 1788
QY 1201 TCTCTGATATCAATCAATGTTCAATTTCAAGCAGATGGCAGCTTCAATGCTTCC 1260
Db 1789 TCTCTGATATCAATCAATGTTCAATTTCAAGCAGATGGCAGCTTCAATGCTTCC 1848
QY 1261 TCTCTCATCAAGAGTGTGTTGAAATCTCAAACTTCACTCATCTGAGTGGAGCAAC 1320
Db 1849 TCTCTCATCAAGAGTGTGTTGAAATCTCAAACTTCACTCATCTGAGTGGAGCAAC 1908

QY 1321 AAGCTATGCCAGTTCTCCCTGTGTCAGAGAACTATGAGAGACTCCGAGAAACAGTCTT 1380
Db 1909 AAGCTATGCCAGTTCTCCCTGTGTCAGAGAACTATGAGAGACTCCGAGAAACAGTCTT 1968
QY 1381 GATTAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGAGACCCGAGGCTTTCAGACAGCAG 1440
Db 1969 GATTAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGAGACCCGAGGCTTTCAGACAGCAG 2028
QY 1441 AGCAAGCATTCATCTGCTGAGAACTCAGACAGAGTGGACCCGCAAGAGTCTT 1500
Db 2029 AGCAAGCATTCATCTGCTGAGAACTCAGACAGAGTGGACCCGCAAGAGTCTT 2088
QY 1501 TCTTCACTGATTCAGAGTGGAGGCTGAGAGCAATTAACACAGCTTCTTTTGGC 1560
Db 2089 TCTTCACTGATTCAGAGTGGAGGCTGAGAGCAATTAACACAGCTTCTTTTGGC 2148
QY 1561 CTTCACAGCCAGCAGCATCTCAGAACTGTGCTGCTGAGCTTAAAGGCTGGAC 1620
Db 2149 CTTCACAGCCAGCAGCATCTCAGAACTGTGCTGAGCTTAAAGGCTGGAC 2208
QY 1621 TCGATATCTTGGCCCCAGACCTTACCCCTGACAGCAGCTGTATTTTGC 1680
Db 2209 TCGATATCTTGGCCCCAGACCTTACCCCTGACAGCAGCTGTATTTTGC 2268
QY 1681 ACAAGTCTCACTTCTGCTGCTGATCTTACAGCAGTTCAGAGGAGTGCATGTTACT 1740
Db 2269 ACAAGTCTCACTTCTGCTGCTGATCTTACAGCAGTTCAGAGGAGTGCATGTTACT 2328
QY 1741 GCTTACAGCTGACGACGCTGCTTGGAGAGCAAGCTTATTTGCTGAGGCGG 1800
Db 2329 GCTTACAGCTGACGACGCTGCTTGGAGAGCAAGCTTATTTGCTGAGGCGG 2388
QY 1801 CAGAAAGCAAGTGAAGAGCTGACCTCGGCGGAGCTGAGTGAAGAGCCCTTTGAA 1860
Db 2389 CAGAAAGCAAGTGAAGAGCTGACCTCGGCGGAGCTGAGTGAAGAGCCCTTTGAA 2448
QY 1861 AAGCAGTTTAAAGCAGAAAGCTGCAATGATTTGAGAGAGCATGATCAAGAGAAC 1920
Db 2449 AAGCAGTTTAAAGCAGAAAGCTGCAATGATTTGAGAGAGCATGATCAAGAGAAC 2508
QY 1921 AGGTACCGGAGAGCTGGGAGAAAGTGGAGAGTCACTTCTTGGGAGCATGAGAA 1980
Db 2509 AGGTACCGGAGAGCTGGGAGAAAGTGGAGAGTCACTTCTTGGGAGCATGAGAA 2568
QY 1981 ATCATTAAGGCTTCC 1995
Db 2569 ATCATTAAGGCTTCC 2583

RESULT 7
US-10-425-114-26234
Sequence 26234, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Ltu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
NUMBER OF SEQ ID NOS: 2003-04-28
SEQ ID NO 26234
LENGTH: 3625
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Clone ID: LIB4119-028-H6_FLI
US-10-425-114-26234

Query Match 99.8%; Score 1991.8; DB 13; Length 3625;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATTGAACTCAATTTGTTACTGAGAGGTTGAGCTGCTGCTGAA 60
DB 692 ATGGCCCATGAGATTGAACTCAATTTGTTACTGAGAGGTTGAGCTGCTGCTGAA 751
QY 61 AGTGAAGCGAAAAAGTGTCTAATTGATAGCCGACATTTGTGAATCAATACATCC 120
DB 752 AGTGAAGCGAAAAAGTGTCTAATTGATAGCCGACATTTGTGAATCAATACATCC 811
QY 121 CACATTTTGAAGCATTAATATCAACTGCTCCCAAGCTTATGAAGCGAAGTTGCAAC 180
DB 812 CACATTTTGAAGCATTAATATCAACTGCTCCCAAGCTTATGAAGCGAAGTTGCAAC 871
QY 181 GACAAAGTTAATTAAGAGCTCATCCAGATTCAGCGAAACATAAGTTGACATTGAT 240
DB 872 GACAAAGTTAATTAAGAGCTCATCCAGATTCAGCGAAACATAAGTTGACATTGAT 931
QY 241 TGCAGTCAGAAAGTTGATTTAGATCAAGCAAGCTCCAGAGATTTGCTCTCTTCA 300
DB 932 TGCAGTCAGAAAGTTGATTTAGATCAAGCAAGCTCCAGAGATTTGCTCTCTTCA 991
QY 301 GACTGTTTCTCACTGACTTCTGAGTAACTGGAAGAAAGCTTCACTCTGCTCCTG 360
DB 992 GACTGTTTCTCACTGACTTCTGAGTAACTGGAAGAAAGCTTCACTCTGCTCCTG 1051
QY 361 CTTCAGAGTGGGTTTGTGAGATTCTCGTTGTTCCCTGAGCCTCTGTGAAGAAATCC 420
DB 1052 CTTCAGAGTGGGTTTGTGAGATTCTCGTTGTTCCCTGAGCCTCTGTGAAGAAATCC 1111
QY 421 ACTCTAGCTCCCTACCTGCAATTTCTCAAGCTTGTCTTACCTGTGCAACATTTGGCC 480
DB 1112 ACTCTAGCTCCCTACCTGCAATTTCTCAAGCTTGTCTTACCTGTGCAACATTTGGCC 1171
QY 481 CGAATTTTCCCAATCTTTTATCTTGGCTGCAAGGAAATGCTTCAACAAAGAGCTGATA 540
DB 1172 CGAATTTTCCCAATCTTTTATCTTGGCTGCAAGGAAATGCTTCAACAAAGAGCTGATA 1231
QY 541 CAGAGAAATGGGATTTGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCCTGACTT 600
DB 1232 CAGAGAAATGGGATTTGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCCTGACTT 1291
QY 601 ATCCCCGAGTCTCAATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAAGAAATTTG 660
DB 1292 ATCCCCGAGTCTCAATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAAGAAATTTG 1351
QY 661 CCGTGTGTTGACAAATAGATGATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTGT 720
DB 1352 CCGTGTGTTGACAAATAGATGATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTGT 1411
QY 721 CTAGTGACTGTTAGCTGGGATCTCCGCTCCGCGCAACATGCTATGCTCATCATG 780
DB 1412 CTAGTGACTGTTAGCTGGGATCTCCGCTCCGCGCAACATGCTATGCTCATCATG 1471
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCCTACT 840
DB 1472 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCCTACT 1531
QY 841 ATATCTTCAAACTTCAATTTTCTGGGCGCAACTCTGAGCTATGAGAAAGATTAAGAC 900
DB 1532 ATATCTTCAAACTTCAATTTTCTGGGCGCAACTCTGAGCTATGAGAAAGATTAAGAC 1591
QY 901 CAGACTGAGATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGAGAAAGCCAAAT 960
DB 1592 CAGACTGAGATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGAGAAAGCCAAAT 1651
QY 961 GAACCTGTCTGTCTCAGAGAGGTGACAGAAAGCGAGAGCCCTCACTCCAGTCCACC 1020
DB 1652 GAACCTGTCTGTCTCAGAGAGGTGACAGAAAGCGAGAGCCCTCACTCCAGTCCACC 1711

QY 1021 TGTGCGGACTCTGCTACTCTCAGAGGACAGAGCAAAAGGCGCGGTGATCCCGCAGCGTG 1080
DB 1712 TGTGCGGACTCTCTACTCTCAGAGGACAGAGCAAAAGGCGCGGTGATCCCGCAGCGTG 1771
QY 1081 CCCAGCGTCCGAGCGTGCAGCGCGCTGCTGTTAGAGACAGCCGCTGATACAGCGCTC 1140
DB 1772 CCCAGCGTCCGAGCGTGCAGCGCGCTGCTGTTAGAGACAGCCGCTGATACAGCGCTC 1831
QY 1141 AGTGGCTGCACTGTTCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCCTTCTTC 1200
DB 1832 AGTGGCTGCACTGTTCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCCTTCTTC 1891
QY 1201 TCTCTGGAATATCAATCAGTTTATATATTCAGCCAGATGAGCAGATCCTTACATGCTTC 1260
DB 1892 TCTCTGGAATATCAATCAGTTTATATATTCAGCCAGATGAGCAGATCCTTACATGCTTC 1951
QY 1261 TCTCTCATCAGAAAGTCTTTGGAATATCAAACTTCCATCTCTGTGATGGAACCAAC 1320
DB 1952 TCTCTCATCAGAAAGTCTTTGGAATATCAAACTTCCATCTCTGTGATGGAACCAAC 2011
QY 1321 AAGCTATGCCAGTTTCTCCCTGTTCAAGAACTTTCGAGAGACTCCCGAAACAGTCTT 1380
DB 2012 AAGCTATGCCAGTTTCTCCCTGTTCAAGAACTTTCGAGAGACTCCCGAAACAGTCTT 2071
QY 1381 GATTAAGGAGAGCGAGCATCCCAAGAGCTGACAGCGGCTTCAAGCAGCAG 1440
DB 2072 GATTAAGGAGAGCGAGCATCCCAAGAGCTGACAGCGGCTTCAAGCAGCAG 2131
QY 1441 AGCAAGCGATTTGATTTGCTGTCAGAACAGCAGAGTGGCAACCGCCAGAGGTCCCTTTA 1500
DB 2132 AGCAAGCGATTTGATTTGCTGTCAGAACAGCAGAGTGGCAACCGCCAGAGGTCCCTTTA 2191
QY 1501 TCTTCATCTGATGAAAGTGGAGCGTGGAGGACATTAACAACAGCTTCTTTTGGC 1560
DB 2192 TCTTCATCTGATGAAAGTGGAGCGTGGAGGACATTAACAACAGCTTCTTTTGGC 2251
QY 1561 CTTCACACGACGAGAGACCTCAGAAAGTCTGAGGCTTGAAGGCTGGCAC 1620
DB 2252 CTTCACACGACGAGAGACCTCAGAAAGTCTGAGGCTTGAAGGCTGGCAC 2311
QY 1621 TCGGATATCTTGGCCCCCGGAGCCTCTACCCCTTCCCTGACAGAGCTGATATTTGCG 1680
DB 2312 TCGGATATCTTGGCCCCCGGAGCCTCTACCCCTTCCCTGACAGAGCTGATATTTGCG 2371
QY 1681 ACAAGTCTCTACACTTCTACTCTGCTCAGCCATCTACGAGGAGCTGCAAGTACTCT 1740
DB 2372 ACAAGTCTCTACACTTCTACTCTGCTCAGCCATCTACGAGGAGCTGCAAGTACTCT 2431
QY 1741 GCCTAAGCTGACGAGGCTGCCACTTGGGAGACCAAGTCTATTCTGCGGCAAGCGG 1800
DB 2432 GCCTAAGCTGACGAGGCTGCCACTTGGGAGACCAAGTCTATTCTGCGGCAAGCGG 2491
QY 1801 CAGAAAGTGAAGAGCTGACTGCGGCGGAGCTGAGCATGAAGAGAGCCCTTTGAA 1860
DB 2492 CAGAAAGTGAAGAGCTGACTGCGGCGGAGCTGAGCATGAAGAGAGCCCTTTGAA 2551
QY 1861 AAGCACTTTAAACGAGAGCTGCCAAATGAAATTTGAGAGAGCATATGTCAAGAAC 1920
DB 2552 AAGCACTTTAAACGAGAGCTGCCAAATGAAATTTGAGAGAGCATATGTCAAGAAC 2611
QY 1921 AGGTCAACGGAAGAGCTGGGGAAGAGGCGAGTCACTTACCTTTCTGGGACAGATGAA 1980
DB 2612 AGGTCAACGGAAGAGCTGGGGAAGAGGCGAGTCACTTACCTTTCTGGGACAGATGAA 2671
QY 1981 ATCATTTAGGTCTCC 1995
DB 2672 ATCATTTAGGTCTCC 2686

RESULT 8
US-10-343-357-17
; Sequence 17, Application US/10343357
; Publication No. US20040058341A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom
 APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
 APPLICANT: VAO, Monique G.; BURFORD, Neil
 APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
 APPLICANT: GANDHI, Ameeta R.; ARVIZU, Chandra S.
 APPLICANT: LEE, Ernestine A.; HAFALIA, April J. A.
 APPLICANT: GRIFFIN, Jennifer A.; TRIBOUTLEY, Catherine M.
 APPLICANT: YUE, Henry; WARREN, Bridget A.
 APPLICANT: NGUYEN, Daniel B.; CHAMLA, Narinder K.
 APPLICANT: KEARNEY, Liam
 TITLE OF INVENTION: PROTEIN PHOSPHATASES
 FILE REFERENCE: PI-0173 PCT
 CURRENT APPLICATION NUMBER: US/10/343,357
 PRIOR APPLICATION NUMBER: PCT/US01/23716
 PRIOR FILING DATE: 2001-07-26
 PRIOR APPLICATION NUMBER: US 60/221,679
 PRIOR FILING DATE: 2000-07-28
 PRIOR APPLICATION NUMBER: US 60/223,272
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: US 60/224,309
 PRIOR FILING DATE: 2000-08-10
 PRIOR APPLICATION NUMBER: US 60/226,728
 PRIOR FILING DATE: 2000-08-18
 PRIOR APPLICATION NUMBER: US 60/229,254
 PRIOR FILING DATE: 2000-09-30
 PRIOR APPLICATION NUMBER: US 60/231,366
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PERL Program
 SEQ ID NO 17
 LENGTH: 3766
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CBI
 US-10-343-357-17

Query Match

Best Local Similarity 99.8%; Score 1991.8; DB 13; Length 3766;
 Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCCCATGAGATGATGGAATCTCAATTTGTAAGAGAGGTTGCTGCTGGAA 60
 DB 538 ATGCCCATGAGATGATGGAATCTCAATTTGTAAGAGAGGTTGCTGCTGGAA 60
 QY 61 AGTGGAAACGGAAGAGTGTCTAATTTAGTACCGGCAATTTGGAAATCAATATCC 120
 DB 598 AGTGGAAACGGAAGAGTGTCTAATTTAGTACCGGCAATTTGGAAATCAATATCC 120
 QY 121 CACATTTTGGAAAGCCATTAATATCACTGCTCCAGCTTATGAGCAAGTTGGCAACG 180
 DB 658 CACATTTTGGAAAGCCATTAATATCACTGCTCCAGCTTATGAGCAAGTTGGCAACG 180
 QY 181 GACAAAGTATTAACAGAGCTCATCCAGATTCAGCGAAACATTAAGTTGACATGAT 240
 DB 718 GACAAAGTATTAACAGAGCTCATCCAGATTCAGCGAAACATTAAGTTGACATGAT 240
 QY 241 TGCAGTGAAGAGTGTGATTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 300
 DB 778 TGCAGTGAAGAGTGTGATTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 300
 QY 301 GACGTGTTTCTCACTGATCTTGGGTAACTGGAGAAAGAGTTCACTGTTCACCTG 360
 DB 838 GACGTGTTTCTCACTGATCTTGGGTAACTGGAGAAAGAGTTCACTGTTCACCTG 360
 QY 361 CTTCAGAGTGGTGTGATGATCTCTCGTTGTTTCCCTGAGCTCTGTAAGAGAAATCC 420
 DB 898 CTTCAGAGTGGTGTGATGATCTCTCGTTGTTTCCCTGAGCTCTGTAAGAGAAATCC 420

QY 421 ACTTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTGGCAACATGGGCAAC 480
 DB 958 ACTTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTGGCAACATGGGCAAC 480
 QY 481 CGAATTTCTCCCAATCTTATATCTGGTCCAGGAGATGCTTCAAGAGAGTGA 540
 DB 1018 CGAATTTCTCCCAATCTTATATCTGGTCCAGGAGATGCTTCAAGAGAGTGA 540
 QY 541 CAGCAAGATGGAGTTGATGATGATTAATGCAAGCTATACCTGTCCAAAGCTGACTT 600
 DB 1078 CAGCAAGATGGAGTTGATGATGATTAATGCAAGCTATACCTGTCCAAAGCTGACTT 600
 QY 601 ATCCCGAGTCTCAATTTCTCGTGTGCTGTGAATGACAGCTTTGAGAAATTTTG 660
 DB 1138 ATCCCGAGTCTCAATTTCTCGTGTGCTGTGAATGACAGCTTTGAGAAATTTTG 660
 QY 661 CCGTGTGAGCAAAATAGTAAATTTCTATGAGAAAGCAAAAGCTCCAAATGATGTT 720
 DB 1198 CCGTGTGAGCAAAATAGTAAATTTCTATGAGAAAGCAAAAGCTCCAAATGATGTT 720
 QY 721 CTAGTCACTGTTTAACTGAGATCTCCGCTCCGACCAATGCTATGCTTACATATG 780
 DB 1258 CTAGTCACTGTTTAACTGAGATCTCCGCTCCGACCAATGCTATGCTTACATATG 780
 QY 781 AAGAGATGAGCATGCTTTAGATGAGTTTACAGATTTGTAAGAAAGAAAGCTACT 840
 DB 1318 AAGAGATGAGCATGCTTTAGATGAGTTTACAGATTTGTAAGAAAGAAAGCTACT 840
 QY 841 ATATCTCCAAATCTCAATTTCTGAGGCACTCTGAGCTATGAGAAAGATTAAGAC 900
 DB 1378 ATATCTCCAAATCTCAATTTCTGAGGCACTCTGAGCTATGAGAAAGATTAAGAC 900
 QY 901 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGACCTGAGAGCAAAAT 960
 DB 1438 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGACCTGAGAGCAAAAT 960
 QY 961 GAACCTGCTCCGCTGCTCTCAGAGGTGAGCAAGAAAGGAGACGCTTCACTCACCC 1020
 DB 1498 GAACCTGCTCCGCTGCTCTCAGAGGTGAGCAAGAAAGGAGACGCTTCACTCACCC 1020
 QY 1021 TGTGCGCATCTGCTTACCTCAGAGGCAAGCAAAAGCCGCTCATCCGCGCAAGC 1080
 DB 1558 TGTGCGCATCTGCTTACCTCAGAGGCAAGCAAAAGCCGCTCATCCGCGCAAGC 1080
 QY 1081 CCGAGGCTGCGCAGCGCTGACGCTGCTGTAGAGCAAGCGCGCTGTACAGGCGCTC 1140
 DB 1618 CCGAGGCTGCGCAGCGCTGACGCTGCTGTAGAGCAAGCGCGCTGTACAGGCGCTC 1140
 QY 1141 AGTGGCTGCACTGTCCGCAAGAGGCTGGAAGCAGCAATTAAGCTCAAGCTTCC 1200
 DB 1678 AGTGGCTGCACTGTCCGCAAGAGGCTGGAAGCAGCAATTAAGCTCAAGCTTCC 1200
 QY 1201 TCTCTGATATCAAAATCAGTTTCAATTTACGCAAGATGAGAGATCTTACATGCTTC 1260
 DB 1738 TCTCTGATATCAAAATCAGTTTCAATTTACGCAAGATGAGAGATCTTACATGCTTC 1260
 QY 1261 TCTCTATCAAGAGATCTTGGAAATCACTCAAACTTCCACTCTGATGAGCAAC 1320
 DB 1798 TCTCTATCAAGAGATCTTGGAAATCACTCAAACTTCCACTCTGATGAGCAAC 1320
 QY 1321 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATGGAAGAGATCTCCGAACGATCCT 1380
 DB 1858 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATGGAAGAGATCTCCGAACGATCCT 1380
 QY 1381 GATTAAGAGAGCAGATCTCCCAAGAGCTGCAAGCGCAAGCTTTCAGACGCAAG 1440
 DB 1918 GATTAAGAGAGCAGATCTCCCAAGAGCTGCAAGCGCAAGCTTTCAGACGCAAG 1440
 QY 1441 AGCAAGGATTTGATTCGCTCAGAAACAGAGCTGCAAGCGCAAGCTTTCAGACGCAAG 1500
 DB 1978 AGCAAGGATTTGATTCGCTCAGAAACAGAGCTGCAAGCGCAAGCTTTCAGACGCAAG 1500
 QY 1501 TCTCACTGATGAAGTGGAGCTGAGAGCAATTTACACACAGCTTCTTTTCCGAC 1560

|||||
Db 2038 TCTCCATGATGGAAGTGGAGGTGAGGACATTAACAACACAGCTTCCTTTGGC 2097
|||
Qy 1561 CTTTCCACCGCCAGCAGCCTCAGAACTGCTGCTGGCTTAAAGGCTGGCAC 1620
|||
Db 2098 CTTTCCACCGCCAGCAGCCTCAGAACTGCTGCTGGCTTAAAGGCTGGCAC 2157
|||
Qy 1621 TCGGATATCTGGGCCCCGAGACCTCTAACCCCTTCCCTGACCAAGAGCTGGTATTTGGC 1680
|||||
Db 2158 TCGGATATCTGGGCCCCGAGACCTCTAACCCCTTCCCTGACCAAGAGCTGGTATTTGGC 2217
|||
Qy 1681 ACAGAGCTCTACACATTTCTACTGCTCAGCCATCTTACGAGGAGCTGGCAGTTACTCT 1740
|||
Db 2218 ACAGAGCTCTACACATTTCTACTGCTCAGCCATCTTACGAGGAGCTGGCAGTTACTCT 2277
|||
Qy 1741 GCCTACAGCTGACCGCAGCTGCCCACTTGGCGAGACCAAGCTTATTTCTGGCGAGCGG 1800
|||
Db 2278 GCCTACAGCTGACCGCAGCTGCCCACTTGGCGAGACCAAGCTTATTTCTGGCGAGCGG 2337
|||
Qy 1801 CAGAAGCAGTGAACAGAGCTGACTGGCGGCGAGCTGGGATGAAGAAGCCCTTTGAA 1860
|||||
Db 2338 CAGAAGCAGTGAACAGAGCTGACTGGCGGCGAGCTGGGATGAAGAAGCCCTTTGAA 2397
|||
Qy 1861 AACGAGTTTAAAGCAGAACTGCAATGAAATTTGAGAGAGCATCATGTCAAGAAC 1920
|||||
Db 2398 AACGAGTTTAAAGCAGAACTGCAATGAAATTTGAGAGAGCATCATGTCAAGAAC 2457
|||
Qy 1921 AGGTACGGAAGAGCTGGGGAAGTGGGCACTGCTTCTTGGCGAGCAGTGA 1980
|||
Db 2458 AGGTACGGAAGAGCTGGGGAAGTGGGCACTGCTTCTTGGCGAGCAGTGA 2517
|||
Qy 1981 ATCATGAGGTCTCC 1995
|||
Db 2518 ATCATGAGGTCTCC 2532

RESULT 9

US-10-648-593-115
; Sequence 115, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 115
; LENGTH: 4790
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-648-593-115

Query Match 99.8%; Score 1991.8; DB 17; Length 4790;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATTGGAACCTCAATTGTTACTGAGAGGTTGGTGGCTTGGCTGGAA 60
Db 184 ATGGCCCATGAGATGATTGGAACCTCAATTGTTACTGAGAGGTTGGTGGCTTGGCTGGAA 243
|||
Qy 61 AGTGAAGCGGAAAGTGTCTAAATTGATAGCGGCGCATTTGGGAAATCAATATACATCC 120
Db 244 AGTGAAGCGGAAAGTGTCTAAATTGATAGCGGCGCATTTGGGAAATCAATATACATCC 303
|||
Qy 121 CACATTTGGAAGCATTAATATCACTGCTCAAGCTTATGAAGCGAGGTTGCAACAG 180
Db 304 CACATTTGGAAGCATTAATATCACTGCTCAAGCTTATGAAGCGAGGTTGCAACAG 363

Qy 181 GACAAAGTTAATTACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
Db 364 GACAAAGTTAATTACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 423
|||
Qy 241 TGCAGTCAGAAAGTTAGTTTACATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 300
Db 424 TGCAGTCAGAAAGTTAGTTTACATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 483
|||
Qy 301 GACTGTTTCTCACTGACTCTTGGGTTAACTGGAGAAAGCTTCAACTGTTCACCTG 360
Db 484 GACTGTTTCTCACTGACTCTTGGGTTAACTGGAGAAAGCTTCAACTGTTCACCTG 543
|||
Qy 361 CTTCGAGGTGGGTTTCTGAGTTCTCTCGTGTGTTCCCTGGGCTCTGTGAAGGAAATCC 420
Db 544 CTTCGAGGTGGGTTTCTGAGTTCTCTCGTGTGTTCCCTGGGCTCTGTGAAGGAAATCC 603
|||
Qy 421 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGTCTACTGTTGCAACATTTGGCCAAAC 480
Db 604 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGTCTACTGTTGCAACATTTGGCCAAAC 663
|||
Qy 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCGACGAGATGTCTTCAACAAAGAGCTGAT 540
Db 664 CGAATTTCTCCCAATCTTTATCTTGGCTGCGACGAGATGTCTTCAACAAAGAGCTGAT 723
|||
Qy 541 CAGCAGAAATGGGATTTGGTTATGTTAATGCGAGCTATACCTGTCCAAAGCCTGACTTT 600
Db 724 CAGCAGAAATGGGATTTGGTTATGTTAATGCGAGCTATACCTGTCCAAAGCCTGACTTT 783
|||
Qy 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 784 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843
|||
Qy 661 CCGTGTGGAACAATCAGTATGATTCTTATGAGAAAGCAAAAGCCTTCAATGATGTTT 720
Db 844 CCGTGTGGAACAATCAGTATGATTCTTATGAGAAAGCAAAAGCCTTCAATGATGTTT 903
|||
Qy 721 CTAGTCACTGTTTACTGGGAACTCCCGCTCCGCAACATGCTATGCTATCATCATG 780
Db 904 CTAGTCACTGTTTACTGGGAACTCCCGCTCCGCAACATGCTATGCTATCATCATG 963
|||
Qy 781 AAGAGATGAGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
Db 964 AAGAGATGAGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1023
|||
Qy 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTTGAATATGAGAAAGATTAAGAAC 900
Db 1024 ATATCTCAAACTTCAATTTTCTGGGCAACTCTTGAATATGAGAAAGATTAAGAAC 1083
|||
Qy 901 CAGACTGAGCATCAGAGGCAAAAGCAAACTCAAGCTCTGCACTGGAGAGCCAAAT 960
Db 1084 CAGACTGAGCATCAGAGGCAAAAGCAAACTCAAGCTCTGCACTGGAGAGCCAAAT 1143
|||
Qy 961 GAACCTGCTCCTGCTCTCAGAGGTTGACAGAAAGCGAGAGCGCCCTCACTCAACCC 1020
Db 1144 GAACCTGCTCCTGCTCTCAGAGGTTGACAGAAAGCGAGAGCGCCCTCACTCAACCC 1203
|||
Qy 1021 TGTGCGACTCTGCTACTCAGAGGCAAGCAAAAGCGCGTGCATCCGCAAGCTG 1080
Db 1204 TGTGCGACTCTGCTACTCAGAGGCAAGCAAAAGCGCGTGCATCCGCAAGCTG 1263
|||
Qy 1081 CCCAGGTCGCCAGCGCTGAGCGCTGTTAGAGAGCAAGCCGCTGGATCAAGGCGCTC 1140
Db 1264 CCCAGGTCGCCAGCGCTGAGCGCTGTTAGAGAGCAAGCCGCTGGATCAAGGCGCTC 1323
|||
Qy 1141 AGTGGCTGCACTGTCTGCGCAGACAGCTGGAAGACAGCAATAGCTCAAGGCTTCTTC 1200
Db 1324 AGTGGCTGCACTGTCTGCGCAGACAGCTGGAAGACAGCAATAGCTCAAGGCTTCTTC 1383
|||
Qy 1201 TCTCTGATATCAATCAAGTTTCAATTTCAAGCCAGATGCGACATCTTACATGCTTC 1260
Db 1384 TCTCTGATATCAATCAAGTTTCAATTTCAAGCCAGATGCGACATCTTACATGCTTC 1443

```

QY 1261 TCCTCATCAGAAAGTCTTGGAAATCTACAAACCTTCCACTACTCTGGATGGACCAAC 1320
DB 1444 TCCTCATCAGAAAGTCTTGGAAATCTACAAACCTTCCACTACTCTGGATGGACCAAC 1503
QY 1321 AAGCTATGCCAGTTCTCCCTGTTCCAGAACTATGGAGCAAGATCTCCGAAACCAAGTCT 1380
DB 1504 AAGCTATGCCAGTTCTCCCTGTTCCAGAACTATGGAGCAAGATCTCCGAAACCAAGTCT 1563
QY 1381 GATTAAGAGAAAGCCAGCAATCCCAAGAGCTGCAGACCCGACAGGCTTTCAGACCAAG 1440
DB 1564 GATTAAGAGAAAGCCAGCAATCCCAAGAGCTGCAGACCCGACAGGCTTTCAGACCAAG 1623
QY 1441 AGCAAGCCATTCATCTGGGTGCAAAACAGCAGCAGTGGCAACCCCAAGAGCTTCTTTA 1500
DB 1624 AGCAAGCCATTCATCTGGGTGCAAAACAGCAGCAGTGGCAACCCCAAGAGCTTCTTTA 1683
QY 1501 TCTCCATGCATTCGAAAGTGGAGGCTGGAGCAATTCACCAACAGCTTCTTTTGGAG 1560
DB 1684 TCTCCATGCATTCGAAAGTGGAGGCTGGAGCAATTCACCAACAGCTTCTTTTGGAG 1743
QY 1561 CTTCACCAAGCCAGCAGCACTTCAGAAAGTCTGGAGCTGGAGCTTAAAGGCTGGAG 1620
DB 1744 CTTCACCAAGCCAGCAGCACTTCAGAAAGTCTGGAGCTGGAGCTTAAAGGCTGGAG 1803
QY 1621 TCGGATATCTGGGCCCCCGAGACTTACCTTCCCTGACCAAGCTGGATATTTGGCC 1680
DB 1804 TCGGATATCTGGGCCCCCGAGACTTACCTTCCCTGACCAAGCTGGATATTTGGCC 1863
QY 1681 ACAGAGCTCTCAACTTCTACTCTGCTCAGACCATCTACGAGGAGGAGGCAAGTACTCT 1740
DB 1864 ACAGAGCTCTCAACTTCTACTCTGCTCAGACCATCTACGAGGAGGAGGCAAGTACTCT 1923
QY 1741 GCTTACAGCTGCAGCAGCTGCGCACTTGGCGGAAACCAAGTCTATTTCTGGCGCAGCGG 1800
DB 1924 GCTTACAGCTGCAGCAGCTGCGCACTTGGCGGAAACCAAGTCTATTTCTGGCGCAGCGG 1983
QY 1801 CAGAAGCCAGTGAAGAGCTGACTCGCGGCGAGCTGGAGTGAAGAGGCCCCCTTTGAA 1860
DB 1984 CAGAAGCCAGTGAAGAGCTGACTCGCGGCGAGCTGGAGTGAAGAGGCCCCCTTTGAA 2043
QY 1861 AAGCAGTTTAAACGAGAAAGCTGCAGAAATTTGGAGAGCACTATGTCAGAGAAC 1920
DB 2044 AAGCAGTTTAAACGAGAAAGCTGCAGAAATTTGGAGAGCACTATGTCAGAGAAC 2103
QY 1921 AGCTCAGCGGAAAGCTGGGAAAGTGGGAGTCACTTCTTTCGCGCAGCATGGAA 1980
DB 2104 AGCTCAGCGGAAAGCTGGGAAAGTGGGAGTCACTTCTTTCGCGCAGCATGGAA 2163
QY 1981 ATCATTGAGTCTCC 1995
DB 2164 ATCATTGAGTCTCC 2178

```

```

RESULT 10
US-10-168-506-2
Sequence 2, Application US/10168506
Publication No. US20040053229A1
GENERAL INFORMATION:
APPLICANT: PLOMAN, GREGORY D.
APPLICANT: MARTINEZ, RICARDO
APPLICANT: WHITE, DAVID
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
APPLICANT: HILL, RON
APPLICANT: PLANAGAN, PETER
TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
FILE REFERENCE: 038602/1351
CURRENT APPLICATION NUMBER: US/10/168, 506
PRIOR FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: PCT/US00/34736
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 2
; LENGTH: 2732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-168-506-2

Query Match      99.8%; Score 1990.2; DB 13; Length 2732;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAATCTCAATTTGTACTAGAGGTTGGTGGCTGTGCTGGAA 60
DB 538 ATGGCCCATGAGATGATGGAATCTCAATTTGTACTAGAGGTTGGTGGCTGTGCTGGAA 597
QY 61 AGTGGACGGAAGAAAGTGGCTTAATGATAGCCGCGCATTTGTGGATATACATATCC 120
DB 598 AGTGGACGGAAGAAAGTGGCTTAATGATAGCCGCGCATTTGTGGATATACATATCC 657
QY 121 CACATTTTGAAGCATTAATATCAATGCTCCAGCTTATGAGAGGAGGTTGCAACAG 180
DB 658 CACATTTTGAAGCATTAATATCAATGCTCCAGCTTATGAGAGGAGGTTGCAACAG 717
QY 181 GACAAATGTTAATTAAGAGCTCAATCCAGCACTTACGGAACATTAAGTGTGACATTGAT 240
DB 718 GACAAATGTTAATTAAGAGCTCAATCCAGCACTTACGGAACATTAAGTGTGACATTGAT 777
QY 241 TGCAGTCAGAAAGTTGATGATTAGCATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
DB 778 TGCAGTCAGAAAGTTGATGATTAGCATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 837
QY 301 GACTGTTTCTCAGTACTTCTGGGTTAACTGAGAAAGCTTCAACTGTCTTCACTG 360
DB 838 GACTGTTTCTCAGTACTTCTGGGTTAACTGAGAAAGCTTCAACTGTCTTCACTG 897
QY 361 CTTCAGAGTGGGTTTGGCTGAATTTCTGCTGTTTCCCTGGCTCTGTGAAGAAATTC 420
DB 898 CTTCAGAGTGGGTTTGGCTGAATTTCTGCTGTTTCCCTGGCTCTGTGAAGAAATTC 957
QY 421 ACTTGAATCTCTACCTGATTTCTCAGCTTTCAGCTTACCTGTTGCCAACTTGGGCAAC 480
DB 958 ACTTGAATCTCTACCTGATTTCTCAGCTTTCAGCTTACCTGTTGCCAACTTGGGCAAC 1017
QY 481 CGAATTCCTCCAAATCTTTATCTTGGCTGCAGAGATGTCCTCAACAAGAGCTGATA 540
DB 1018 CGAATTCCTCCAAATCTTTATCTTGGCTGCAGAGATGTCCTCAACAAGAGCTGATA 1077
QY 541 CAGCAGATGGAGTGGTATATGTTAAATGCGACATTAACCTGTCGAAGCTGACTTT 600
DB 1078 CAGCAGATGGAGTGGTATATGTTAAATGCGACATTAACCTGTCGAAGCTGACTTT 1137
QY 601 ATCCCGAGTCTGATTTCTGGGTGCTGCTGTGAATGACAGCTTTTGGAGAAATTTTG 660
DB 1138 ATCCCGAGTCTGATTTCTGGGTGCTGCTGTGAATGACAGCTTTTGGAGAAATTTTG 1197
QY 661 CGGTGTTGCAAAATCAGTATGATTTCAATGAGAAACCAAGCTCCATGATGTT 720
DB 1198 CGGTGTTGCAAAATCAGTATGATTTCAATGAGAAACCAAGCTCCATGATGTT 1257
QY 721 CTAGTGAATGTTAGCTGGGATCTCCCGCTCCGCAACATGCTATCGCTACATCATG 780
DB 1258 CTAGTGAATGTTAGCTGGGATCTCCCGCTCCGCAACATGCTATCGCTACATCATG 1317
QY 781 AAGAGATGAGCATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
DB 1318 AAGAGATGAGCATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1377
QY 841 ATATCTCCAAATCTCAATTTCTGGGCAACTCTGCACTATGAGAAAGATTTAAGAAC 900
DB 1378 ATATCTCCAAATCTCAATTTCTGGGCAACTCTGCACTATGAGAAAGATTTAAGAAC 1437
QY 901 CAGACTGAGACATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGGAGAGCCAAAT 960
DB 1438 CAGACTGAGACATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGGAGAGCCAAAT 1497

```

QY 961 GAACCTGTCCCTGCTGCTCTCAAGAGGTGGA CAGAAAGCGAGACGCCCTCACTGTCACCC 1020
DB 1498 GAACCTGTCCCTGCTGCTCTCAAGAGGTGGA CAGAAAGCGAGACGCCCTCACTGTCACCC 1557
QY 1021 TGTGCCACTCTGCTCACTCAAGAGGTGGA CAGAAAGCGAGACGCCCTCACTGTCACCC 1080
DB 1558 TGTGCCACTCTGCTCACTCAAGAGGTGGA CAGAAAGCGAGACGCCCTCACTGTCACCC 1617
QY 1081 CCCAGCGTCCAGCGTGCAGCGCTGCTGTTAAGAGACAGCCCGTGTACAGAGCGCTC 1140
DB 1618 CCCAGCGTCCAGCGTGCAGCGCTGCTGTTAAGAGACAGCCCGTGTACAGAGCGCTC 1677
QY 1141 AGTGGGCTGACCTGCTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCGCTTCC 1200
DB 1678 AGTGGGCTGACCTGCTCCGAGAGAGGCTGGAAGACAGCAATAGCTCAAGCGCTTCC 1737
QY 1201 TCTCTGATATCAAAATGATTTTCAATATTCAGCCAGCATGAGCATCTTACATGCTTC 1260
DB 1738 TCTCTGATATCAAAATGATTTTCAATATTCAGCCAGCATGAGCATCTTACATGCTTC 1797
QY 1261 TCTCTCATCAAGATGCTTTGGAATATCTACAAACCTTCACTA CTCTGATGAGGACCAAC 1320
DB 1798 TCTCTCATCAAGATGCTTTGGAATATCTACAAACCTTCACTA CTCTGATGAGGACCAAC 1857
QY 1321 AAGCTATGCAAGTTCCTCCCTGTTCAAGAACTATCGAGAGACATCCCGAAACAGTCTC 1380
DB 1858 AAGCTATGCAAGTTCCTCCCTGTTCAAGAACTATCGAGAGACATCCCGAAACAGTCTC 1917
QY 1381 GATTAAGAGAGAGCAGCATCCCAAGAGCTGAGACCGCAGCGCTTCAAGAGCAG 1440
DB 1918 GATTAAGAGAGAGCAGCATCCCAAGAGCTGAGACCGCAGCGCTTCAAGAGCAG 1977
QY 1441 AGCAAGGATGATTCGATTCAGAAACAGAGACAGTGGCA CCGCCAGAGCTCCCTTTA 1500
DB 1978 AGCAAGGATGATTCGATTCAGAAACAGAGACAGTGGCA CCGCCAGAGCTCCCTTTA 2037
QY 1501 TCTTCACTGATCGAAGTGGAGAGCTGAGAGACAACTTACACACAGCTTCCTTTGCGC 1560
DB 2038 TCTTCACTGATCGAAGTGGAGAGCTGAGAGACAACTTACACACAGCTTCCTTTGCGC 2097
QY 1561 CTTTCCACAGCAGCAGCAGCATCAAGAGTGTGCTGAGGAGCTTAAAGGCTGGCAC 1620
DB 2098 CTTTCCACAGCAGCAGCAGCATCAAGAGTGTGCTGAGGAGCTTAAAGGCTGGCAC 2157
QY 1621 TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGTATTTGCCC 1680
DB 2158 TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGTATTTGCCC 2217
QY 1681 ACAGAGTCTTCACTTCTGCTGCTCAAGCATCTTACGAGAGCAGTGCAGATTA CTCT 1740
DB 2218 ACAGAGTCTTCACTTCTGCTGCTCAAGCATCTTACGAGAGCAGTGCAGATTA CTCT 2277
QY 1741 GCGTACAGCTGACAGCAGCTGCCACTTGGGAGAGACCAAGCTTATTCTGTCGCGAGCGG 1800
DB 2278 GCGTACAGCTGACAGCAGCTGCCACTTGGGAGAGACCAAGCTTATTCTGTCGCGAGCGG 2337
QY 1801 CAGAAAGCAGTGA CAGAGCTGATCGCGGCGAGAGCTGAGCATGAAGAGACCCCTTTGAA 1860
DB 2338 CAGAAAGCAGTGA CAGAGCTGATCGCGGCGAGAGCTGAGCATGAAGAGACCCCTTTGAA 2397
QY 1861 AAGCAGTTTAAAGCAGAGCTGCGCAATGGAATTTGAGAGAGCATCATGTCAAGAAC 1920
DB 2398 AAGCAGTTTAAAGCAGAGCTGCGCAATGGAATTTGAGAGAGCATCATGTCAAGAAC 2457
QY 1921 AAGTCAAGGAGAGAGCTGGGAGAGAGTGGAGTCAAGTCTTATTTGCGGAGCATGGA 1980
DB 2458 AAGTCAAGGAGAGAGCTGGGAGAGAGTGGAGTCAAGTCTTATTTGCGGAGCATGGA 2517
QY 1981 ATCATTTGAGGTCTCC 1995
DB 2518 ATCATTTGAGGTCTCC 2532

RESULT 11
US-10-094-749-673
; Sequence 673, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YUKI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKI, ICHIRO
; APPLICANT: SEKI, NAOTAKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094, 749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350, 435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; LENGTH: 2102
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-673
Query Match 99.7%; Score 1988.6; DB 16; Length 2102;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGTGAATTGGAACCTCAATTTTCTGAGAGTGTGCTGCTGGA 60
DB 56 ATGGCCCATGAGTGAATTGGAACCTCAATTTTCTGAGAGTGTGCTGCTGGA 115
QY 61 AGTGAACCGAAAGAGTGTGCTGAATTGATAGCCGCGCAATTTGTGAATACATATCC 120
DB 116 AGTGAACCGAAAGAGTGTGCTGAATTGATAGCCGCGCAATTTGTGAATACATATCC 175
QY 121 CACATTTGGAAGCATTAATATCACTGCTCAAGCTTATGAGAGGAGTTGCAACAG 180
DB 176 CACATTTGGAAGCATTAATATCACTGCTCAAGCTTATGAGAGGAGTTGCAACAG 235
QY 181 GACAAAGTGTAAATTA CAGAGCTCATCCAGCATTCAGCAACATTAAGGTTGACATTGAT 240
DB 236 GACAAAGTGTAAATTA CAGAGCTCATCCAGCATTCAGCAACATTAAGGTTGACATTGAT 295
QY 241 TGCAGTCAAGAGTGTAAATTA CAGAGCTCATCCAGCATTCAGCAACATTAAGGTTGACATTGAT 300
DB 296 TGCAGTCAAGAGTGTAAATTA CAGAGCTCATCCAGCATTCAGCAACATTAAGGTTGACATTGAT 355
QY 301 GACTGTTTCTCACTGTACTTCTGAGTAACTGAGAGAGCTTCAACTCTGTTCACTG 360
DB 356 GACTGTTTCTCACTGTACTTCTGAGTAACTGAGAGAGCTTCAACTCTGTTCACTG 415
QY 361 CTTCAGAGTGGTGTGCTGAGTCTCTGTTGTTTCCCTGAGCCTCTGTAAGGAAATCC 420
DB 416 CTTCAGAGTGGTGTGCTGAGTCTCTGTTGTTTCCCTGAGCCTCTGTAAGGAAATCC 475
QY 421 ACTTAAGTCCCTAATCTGCAATTTCTCAAGCTTGTGTTA CTTGTTGCAACATTTGGGCCAAC 480

Db 476 ACTAGTCCCTACCTGCAATTTCTAGCCCTGCTTACCTGTGCAACATTTGGGCGAAC 535
Qy 481 CGAATTCCTCCCAATTTTATTTGCTGCGAGCAGATGTCCTCAACAGAGGCTGATA 540
Db 536 CTATATCTTCCCAATCTTATCTTGGCTGCCGAGATGTCCTCAACAGAGGCTGATA 540
Qy 541 CAGCAGAAATGGGATGTTAGTGTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 595
Db 596 CAGCAGAAATGGGATGTTAGTGTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 600
Qy 601 ATCCCGAGTCTCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 655
Db 656 ATCCCGAGTCTCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 CCGTGTGTTGCAAAATCAGTATGATTTGATTTGAAAGAAAAGCTTCCATGATGTGT 715
Db 716 CCGTGTGTTGCAAAATCAGTATGATTTGATTTGAAAGAAAAGCTTCCATGATGTGT 720
Qy 721 CTATGCACTGTGTTAGTGTGAGATCTCCGCTCCGCAACATGCTATGCTATGATGATG 775
Db 776 CTATGCACTGTGTTAGTGTGAGATCTCCGCTCCGCAACATGCTATGCTATGATGATG 780
Qy 781 AAGAGATGAGACATGCTTTAGATGAGATGCTTACAGATTTGAAAGAAAAGCTTACT 835
Db 836 AAGAGATGAGACATGCTTTAGATGAGATGCTTACAGATTTGAAAGAAAAGCTTACT 840
Qy 841 ATATCTCAAACTTCAATTTTCTGGCCAACTCTCTGCACTATGAGAGAAATTTAAGAC 895
Db 896 ATATCTCAAACTTCAATTTTCTGGCCAACTCTCTGCACTATGAGAGAAATTTAAGAC 900
Qy 901 CAGATCTGAGATCAAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTGAGAGAGCCAAAT 955
Db 956 CAGATCTGAGATCAAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTGAGAGAGCCAAAT 960
Qy 961 GAACCTGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1015
Db 1016 GAACCTGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Qy 1021 TGTGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075
Db 1076 TGTGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Qy 1081 CCCAGCGTCCAGCGTGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1135
Db 1136 CCCAGCGTCCAGCGTGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Qy 1141 AGTGGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1195
Db 1196 AGTGGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Qy 1201 TCTCTGATATCAATCATGTTTCAATTTCAAGCAGCATGCGATCTTCAATGCTTC 1255
Db 1256 TCTCTGATATCAATCATGTTTCAATTTCAAGCAGCATGCGATCTTCAATGCTTC 1260
Qy 1261 TCTCTGATATCAATCATGTTTCAATTTCAAGCAGCATGCGATCTTCAATGCTTC 1315
Db 1316 TCTCTATCAAGAAAGATGCTTGAATACTCAAACTTCACTGATGATGAGCAAC 1320
Qy 1321 AAGCTATGCAATGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1375
Db 1376 AAGCTATGCAATGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Qy 1381 GATTAAGAGAAAGCAGATCTCCCAAGAACTGCAAGCTGCGAGGCTTCAACAGCGAG 1435
Db 1436 GATTAAGAGAAAGCAGATCTCCCAAGAACTGCAAGCTGCGAGGCTTCAACAGCGAG 1440
Qy 1441 AGCAAGCGATTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1495
Db 1496 AGCAAGCGATTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Qy 1501 TCTCACTGATGAGAGTGGAGCGTGGAGCAATTTCAACAGAGCTTCTTTTGCGC 1555
Db 1556 TCTCACTGATGAGAGTGGAGCGTGGAGCAATTTCAACAGAGCTTCTTTTGCGC 1560

Db 1556 TCTCACTGATGAGAGTGGAGCGTGGAGCAATTTCAACAGAGCTTCTTTTGCGC 1615
Qy 1561 CTTTCCACAGCAGCAGACCTTCAAGAGTGTGCTGCGCTGGGCTTTAAGGGCTGGAC 1620
Db 1616 CTTTCCACAGCAGCAGACCTTCAAGAGTGTGCTGCGCTGGGCTTTAAGGGCTGGAC 1675
Qy 1621 TCGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACACAGAGCTGATATTTGCG 1680
Db 1676 TCGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACACAGAGCTGATATTTGCG 1735
Qy 1681 ACAGAGTCTCAACCTTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1736 ACAGAGTCTCAACCTTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1795
Qy 1741 GCTTACAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Db 1796 GCTTACAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1855
Qy 1801 CAGAAAGCAAGTACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
Db 1856 CAGAAAGCAAGTACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1915
Qy 1861 AAGCATTTTAAACGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
Db 1916 AAGCATTTTAAACGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1975
Qy 1921 AGGTCAAGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Db 1976 AGGTCAAGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2035
Qy 1981 ATCATTTAGGCTTCC 1995
Db 2036 ATCATTTAGGCTTCC 2050

RESULT 12
US-10-296-115-520
/ Sequence 520, Application US/10296115
/ Publication No. US20040053248A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyseng Inc
/ TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
/ FILE REFERENCE: 784PCT
/ CURRENT APPLICATION NUMBER: US/10/296,115
/ PRIOR FILING DATE: 2002-11-18
/ PRIOR APPLICATION NUMBER: US09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ PRIOR APPLICATION NUMBER: US09/552,317
/ NUMBER OF SEQ ID NOS: 1478
/ SEQ ID NO 520
/ LENGTH: 2966
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(2966)
/ OTHER INFORMATION: n = a, t, c or g
US-10-296-115-520

Query Match 99.7%; Score 1988.6; DB 13; Length 2966;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATTTGAATCTCAATTTGTTACTGAGAGGTGGTGGCTCTGCTGAA 60
Db 23 ATGGCCCATGAGATGATTTGAATCTCAATTTGTTACTGAGAGGTGGTGGCTCTGCTGAA 82
Qy 61 AGTGAAGGAGAAAAAGTGTGCTGCTAATTTGATTTGATTTGCTGCTGCTGCTGCTGCTG 120
Db 83 AGTGAAGGAGAAAAAGTGTGCTGCTAATTTGATTTGATTTGCTGCTGCTGCTGCTGCTG 142
Qy 121 CACATTTTGAAGCCATTATATCAACTGCTTCAAGCTTATGAGAGCAAGTTGCAACAG 180

Db	143	CACTTTGGAGCATTAATCACTGCTCCAAAGCTTAAGAGGAAAGGTTGCAAC	202
Qy	181	GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTTGAT	240
Db	203	GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTTGAT	262
Qy	241	TGCAAGTGAAGGTTAGTTAGATCAAGGCTCCAGAGATGTTGCTCTCTCTCA	300
Db	263	TGCAAGTGAAGGTTAGTTAGATCAAGGCTCCAGAGATGTTGCTCTCTCTCA	322
Qy	301	GACTGTTTCTCACTGTACTTCTGGGTAACTGGAGAGAGCTTCACTCTGTCACTG	360
Db	323	GACTGTTTCTCACTGTACTTCTGGGTAACTGGAGAGAGCTTCACTCTGTCACTG	382
Qy	361	CTTGACAGTGGGTTTCTGTAGATTCTCTGTTTCTCTGCTCTGTGAGAGAAATCC	420
Db	383	CTTGACAGTGGGTTTCTGTAGATTCTCTGTTTCTCTGCTCTGTGAGAGAAATCC	442
Qy	421	ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTCTTACCTGTGCAATTTGGGCAAC	480
Db	443	ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTCTTACCTGTGCAATTTGGGCAAC	502
Qy	481	CGAATTCCTCCCAATCTTTATCTTGTGCTGCGAGAGATGCTCTCAACAAGAGCTGATA	540
Db	503	CGAATTCCTCCCAATCTTTATCTTGTGCTGCGAGAGATGCTCTCAACAAGAGCTGATA	562
Qy	541	CAGCAGAAATGGGATGGTTATGTGTTAAATGCGAGCTATACCTGTCCAAAGCTGACTTT	600
Db	563	CAGCAGAAATGGGATGGTTATGTGTTAAATGCGAGCTATACCTGTCCAAAGCTGACTTT	622
Qy	601	ATCCCGAGTCTAATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG	660
Db	623	ATCCCGAGTCTAATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG	682
Qy	661	CCGTGTTTGAACAATCAGTGAATTTCTATGAGAAACAAAGCTTCAATGATGTT	720
Db	683	CCGTGTTTGAACAATCAGTGAATTTCTATGAGAAACAAAGCTTCAATGATGTT	742
Qy	721	CTATGTCAGCTGTTTAACTGGGATCTCCGCTCCGCACTCGCTATGCTTACATCATG	780
Db	743	CTATGTCAGCTGTTTAACTGGGATCTCCGCTCCGCACTCGCTATGCTTACATCATG	802
Qy	781	AAGAAGATGGAATGCTTTTAAATGAGCTTACAGATTTGTGAAAGAAACCTTACT	840
Db	803	AAGAAGATGGAATGCTTTTAAATGAGCTTACAGATTTGTGAAAGAAACCTTACT	862
Qy	841	ATATCTCCAATCTTCAATTTTCTGGGCACTCTGTGACTATGAGAGATTTAAGAAC	900
Db	863	ATATCTCCAATCTTCAATTTTCTGGGCACTCTGTGACTATGAGAGATTTAAGAAC	922
Qy	901	CAGACTGAGCATCAAGGCGCAAAAGCAAACTCAAGCTGTGCACTGAGAAACCAAT	960
Db	923	CAGACTGAGCATCAAGGCGCAAAAGCAAACTCAAGCTGTGCACTGAGAAACCAAT	982
Qy	961	GAACTGTCCTCTGCTGTCTAGAGAGGTGACAAAAAGCGCAAGCCCTCATCTCCACC	1020
Db	983	GAACTGTCCTCTGCTGTCTAGAGAGGTGACAAAAAGCGCAAGCCCTCATCTCCACC	1042
Qy	1021	TGTGCGACTCTGTCTACTCAGAGAGCAGAGCAAAAGCCGCTGATCCCGCAGAGCTG	1080
Db	1043	TGTGCGACTCTGTCTACTCAGAGAGCAGAGCAAAAGCCGCTGATCCCGCAGAGCTG	1102
Qy	1081	CCCAAGGCTGCCAGCGTGCAGCGCTGTGTTAGAGAGAGCCCGCTGTATCAGGCGCTC	1140
Db	1103	CCCAAGGCTGCCAGCGTGCAGCGCTGTGTTAGAGAGAGCCCGCTGTATCAGGCGCTC	1162
Qy	1141	AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCGCTTCTTC	1200
Db	1163	AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCGCTTCTTC	1222
Qy	1201	TCTCTGATATCAAAATCAGTTTCAATTTCAAGCAGCATGCGACATCTTACATGCTTC	1260

Db	1223	TCTCTGATATCAAAATCAGTTTCAATTTCAAGCAGCATGCGACATCTTACATGCTTC	1282
Qy	1261	TCTCTGATATCAAAATCAGTTTCAATTTCAAGCAGCATGCGACATCTTACATGCTTC	1320
Db	1283	TCTCTGATATCAAAATCAGTTTCAATTTCAAGCAGCATGCGACATCTTACATGCTTC	1342
Qy	1321	AGCTATGCGAGTTCTCCCTGTTCAGGAACATATCGAGAGAGACTCCGAAACAGCTCT	1380
Db	1343	AGCTATGCGAGTTCTCCCTGTTCAGGAACATATCGAGAGAGACTCCGAAACAGCTCT	1402
Qy	1381	GATTAAGAGAGAGCAGCATCTCCCAAGAACTGCAAGCCGACAGCTTTCAGACGCAAG	1440
Db	1403	GATTAAGAGAGAGCAGCATCTCCCAAGAACTGCAAGCCGACAGCTTTCAGACGCAAG	1462
Qy	1441	AGCAAGCATTTGATTTCTGCTCAGAACAGAGAGAGTGCACCCGCCAGAGTCCCTTTTA	1500
Db	1463	AGCAAGCATTTGATTTCTGCTCAGAACAGAGAGAGTGCACCCGCCAGAGTCCCTTTTA	1522
Qy	1501	TCTCACTGATGAGAGTGGAGAGCTGAGAGACAAATTAACAACAAGCTTCTTTTCGCG	1560
Db	1523	TCTCACTGATGAGAGTGGAGAGCTGAGAGACAAATTAACAACAAGCTTCTTTTCGCG	1582
Qy	1561	CTTTTCAACAGCCAGAGACACTCAGAAATCTGTGCTGCTGGGCTTTAAGGCTGGCAC	1620
Db	1583	CTTTTCAACAGCCAGAGACACTCAGAAATCTGTGCTGCTGGGCTTTAAGGCTGGCAC	1642
Qy	1621	TGGGATATCTTGGCCCCCAGACCTTACCCCTTCTCTGACACAGAGCTGTATTTTGGC	1680
Db	1643	TGGGATATCTTGGCCCCCAGACCTTACCCCTTCTCTGACACAGAGCTGTATTTTGGC	1702
Qy	1681	ACAGAGTCTCAACTTTACTCTGCTGCTGAGCATGACGAGGAGAGTGCAGTTACTCT	1740
Db	1703	ACAGAGTCTCAACTTTACTCTGCTGCTGAGCATGACGAGGAGAGTGCAGTTACTCT	1762
Qy	1741	GCCTTACAGCTGACAGCAGCTGCCACTTGGGAGAGCAAACTTATCTGTGCGCAGGCGG	1800
Db	1763	GCCTTACAGCTGACAGCAGCTGCCACTTGGGAGAGCAAACTTATCTGTGCGCAGGCGG	1822
Qy	1801	CAGAACCAAGTACAGAGCTGACTGCGCGGAGAGCTGCGCATTAAGAGAGCCCTTTGAA	1860
Db	1823	CAGAACCAAGTACAGAGCTGACTGCGCGGAGAGCTGCGCATTAAGAGAGCCCTTTGAA	1882
Qy	1861	AAGCAGTTTAAAGCGAAGCTGCGCAATGGAATTTGAGAGAGCATCATGTGAGAAAC	1920
Db	1883	AAGCAGTTTAAAGCGAAGCTGCGCAATGGAATTTGAGAGAGCATCATGTGAGAAAC	1942
Qy	1921	AGGTCAAGGAGAGAGCTGGGAGAAAGTGGCAGTCACTTATGCTTTTCGGGACAGATGAA	1980
Db	1943	AGGTCAAGGAGAGAGCTGGGAGAAAGTGGCAGTCACTTATGCTTTTCGGGACAGATGAA	2002
Qy	1981	ATCATTGAGGTCTCC	1995
Db	2003	ATCATTGAGGTCTCC	2017

RESULT 13
 US-10-072-012-257
 ; Sequence 257, Application US/10072012
 ; Publication No. US2004003493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Zertusen, Bryan
 ; APPLICANT: Paturajan, Meera
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Li, Li
 ; APPLICANT: Gangoli, Bha
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Anderson, David W.
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Taupier Jr, Raymond J.

Query Match	98.4%;	Score 1962.2;	DB 13;	Length 2071;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 1983:	Conservation			

Conservative	Mismatches	Indels	Gaps
0	3	9	1

1 AAGGCCAAGAGATGATGGAACTCAAATGTTACTGAGAGGTTGTCGCTGCTGGA 60
|||
2 |||
3 |||

61 AAGGCCCATGAGATGATTTGGAACCAATTGTTACTGAGAGTTGTGGCTCTGCTGGAA 120

61 AGTGAACGAAAAAGTCTCTAATTGATAGCCGGCAATTGTGAATACATACATCC 120

121 AGTGACGGAAAAAGTCTCTAATTGATTAGCCGGCCATTGTGGATACAATCAATG 100

121 CACATTTGGAGCCATTATAATCACTGCTCCAAGCTTATGAAGCGAAGCTTCATTA
|||||.....CCTTAACTACATCC 180

181 CACATTTTGAAGCCATTATAITCACTGCTCCAAGCTTTCGACCGGGCAGCAGGACG
180

181 GACAACTGTTAATTACAGAGCTCATCCAGCAATTCAGCAG
.....TTCAGGCCGAAGTGACACAG 240

241 GACAAAGTCTTAATTACAGACTGCGGGCCTTTTCATGC
240

241 TGCAGTCAGAAGCATTGCGA GCTTGGACATTTGAT 300

301 TTGACGCTACACACCTGGCTTCTCTCTCTTCA 300

301 CCGGAGGTTACGATCAAAGCTCCAGATGTGGCTCTCTCTTCA 360

501 GACGGTTTCTCACTGTACTCTGGTAACTGGAGAGACTTCAACTCTGTACCTG 360

561 GACAGCTTTTCTCACTGTACTTCTGGTAACTGGAGAGAGCTTCAACTCTGTTCACTG 420

361 CTTCACGCTGGCTTCGTGAGTTCCTCGTGTTCCTCCGACCTCTGTGAAGAAATCC 420

|||||TCTGTTGGCCCTGTGAAGGAAATCC 420

421 CTGGAGGTGGGTTTGGTGAATCTCTCGTTGTTTCCCTGGCCCTGTGAAAGAAAATCC 480

421 ACTCTAGTCCCTACCTGGATTTCTCAGCCTTGCTACTGTGTCGCAATTTGGGCCAAC 480

481 ACTGAGTCCCTACCTGATTTCTCAGCCTTGCTACTGTGTCGCAATTTGGGCCAAC 540

481 CGAATCTCTCCAAATCTTTATCTTGGCTGTCAGCGATGCTGCTCAACAGAGCTGATA 540

541 CGAATCTCTCCAAATCTTTATCTTGGCTGTCAGCGATGCTGCTCAACAGAGCTGATA 540

541 CAGAGAAATGGGATTGGTTATGTGTAAATGCGAGCTATACCTGTGCCAAAGCTGATG 600

601 CAGAGAAATGGGATTGGTTATGTGTAAATGCGAGCTATACCTGTGCCAAAGCTGATG 600

601 ATCCCGAGTCTCATTTCTGCGGTGCTGGATGACAGTTTGTGTGAAAAATTTTG 660

661 ATCCCGAGTCTCATTTCTGCGGTGCTGGATGACAGTTTGTGTGAAAAATTTTG 720

661 CCGTGGTTGGAAATCAGTAGATTTCAATGAAAGCAAAGGCTCCAAATGATGTGT 720

721 CCGTGGTTGGAAATCAGTAGATTTCAATGAAAGCAAAGGCTCCAAATGATGTGT 780

721 CTATGCACTGTTTAGCTGGGATCTCCCGCTCCGCGCAATCGTTATGCTTACATCATG 780

781 CTATGCACTGTTTAGCTGGGATCTCCCGCTCCGCGCAATCGTTATGCTTACATCATG 840

781 AAGAGAGAGGACATGCTTTATGATGAAGCTTACAGATTTGTGAAAAAAAAGACTACT 840

841 AAGAGAGAGGACATGCTTTATGATGAAGCTTACAGATTTGTGAAAAAAAAGACTACT 900

841 ATATCTCAAACTTCAATTTCTGGGCGCAACTCTGGAATGAGAAAGATTTAAGAAC 900

901 ATATCTCAAACTTCAATTTCTGGGCGCAACTCTGGAATGAGAAAGATTTAAGAAC 960

901 CAGACTGAGCATCAGGGCCAAAGACAACTCAAGCTGCTCACTGAGAGAGGCAAT 960

961 CAGACTGAGCATCAGGGCCAAAGACAACTCAAGCTGCTCACTGAGAGAGGCAAT 1020

961 GAACCTGTCCCTGCTGCTCTCAGAGGGTGGACGAAGAAGCGAGGCCCTCAATCCACC 1020

1021 GAACCTGTCCCTGCTGCTCTCAGAGGGTGGACGAAGAAGCGAGGCCCTCAATCCACC 1080

1021 TGTGCGCACTGCTACCTCAGAGGCGACAGACAAAGGCCCTGTCATCCGCGCAGAGGTG 1080

1081 TGTGCGCACTGCTACCTCAGAGGCGACAGACAAAGGCCCTGTCATCCGCGCAGAGGTG 1132

1081 CCCAGCGTCCCAAGCTGACAGCGCTGCTGTTAAGAGCAGCCCGCTGTGACAGCGCTC 1140

1133 -CCAGCGTCCCAAGCTGACAGCGCTGCTGTTAAGAGCAGCCCGCTGTGACAGCGCTC 1140

1141 AATGGGCTGCACTGTCCGCGAGACAGCGCTGGAAGACAGCAATAAGCTCAAGCGTCTTC 1200

1192 AATGGGCTGCACTGTCCGCGAGACAGCGCTGGAAGACAGCAATAAGCTCAAGCGTCTTC 1251

1201 TCTCTGGAATCAAAATCAGTTTCAATATCAGCCAGATGAGCATCTTACATGAGCTTC 1260

1252 TCTCTGGAATCAAAATCAGTTTCAATATCAGCCAGATGAGCATCTTACATGAGCTTC 1311

1261 TCTCTCATCAGAAATGCTTTGGAAATCTCAAACTTCCATCTCTGTGATGGAGCAAC 1320

1312 TCTCTCATCAGAAATGCTTTGGAAATCTCAAACTTCCATCTCTGTGATGGAGCAAC 1371

1321 AAGCTATGCGAATTTCCCTCTGTTCAAGAACTATGGAAGCAGCTCCGAAACCAAGTCTT 1380

1372 AAGCTATGCGAATTTCCCTCTGTTCAAGAACTATGGAAGCAGCTCCGAAACCAAGTCTT 1431

1381 GATTAAGAGAGAGCAGCATCCCAAGAAAGCTGCGAGCCGAGGCTTCAAGACGCCAG 1440

1432 GATTAAGAGAGAGCAGCATCCCAAGAAAGCTGCGAGCCGAGGCTTCAAGACGCCAG 1491

1441 ACGAAGCATGGATGGGTGAGAACCAAGCAGCAGTGGCAACCGCCGAGAGTCCCTTTTA 1500

Qy	1501	TCTCCACGTGCAATCGAAGTGGGAGCCTGGAGGACATTAACCAACAGCTTCCTTTGCGC	1560
Db	1552	TCTCCACGTGCAATCGAAGTGGGAGCCTGGAGGACATTAACCAACAGCTTCCTTTGCGC	1611
Qy	1561	CTTTCACACGACGACGACGACCTCTACGAAGTCTGTGGCCTTGGCCTTAAAGGCTGGCAC	1620
Db	1612	CTTTCACACGACGACGACGACCTCTACGAAGTCTGTGGCCTTGGCCTTAAAGGCTGGCAC	1671
Qy	1621	TCGGATATCTTGGCCCCCGACAGCCTCTACCCCTTCCTGACACGAGCTGGATATTTGGC	1680
Db	1672	TCGGATATCTTGGCCCCCGACAGCCTCTACCCCTTCCTGACACGAGCTGGATATTTGGC	1733
Qy	1681	ACAAGTCTCTACACTTCTACTCTTGCTCTGAGCCATCTACGAGGACGTCGCCAGTTATCTT	1740
Db	1732	ACAAGTCTCTACACTTCTACTCTTGCTCTGAGCCATCTACGAGGACGTCGCCAGTTATCTT	1791
Qy	1741	GCTTACAGCTGCAGCCTGAGCTGCCACTTGGCGGAGACCAAGTCTATTTCTGTGCGACGGCGG	1800
Db	1792	GCTTACAGCTGCAGCCTGAGCTGCCACTTGGCGGAGACCAAGTCTATTTCTGTGCGACGGCGG	1851
Qy	1801	CAGAAGCCAAGTGA CAGAGCTGACTCGCGGCGGAGCTGSCATGAAGAGAGCCCTTTGAA	1860
Db	1852	CAGAAGCCAAGTGA CAGAGCTGACTCGCGGCGGAGCTGSCATGAAGAGAGCCCTTTGAA	1911
Qy	1861	AAGCAGTTTAAACGAGAGAGCTGCCAATTTGGAATTTGGAGAGCATATTCAGAGAAC	1920
Db	1912	AAGCAGTTTAAACGAGAGAGCTGCCAATTTGGAATTTGGAGAGCATATTCAGAGAAC	1971
Qy	1921	AGGTCAACGGGAAAGAGCTGGGGAAAGATGGGACATGCTAGCTTTTCGGGCGAGCATGAA	1980
Db	1972	AGGTCAACGGGAAAGAGCTGGGGAAAGATGGGACATGCTAGCTTTTCGGGCGAGCATGAA	2031
Qy	1981	ATCATTTAGGTCTCC 1995	
Db	2032	ATCATTTAGGTCTCC 2046	

RESULT 14
US-10-072-012-255
Sequence 255, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: zernhusen, Bryan
APPLICANT: Patrușajan, Meera
APPLICANT: Shimbets, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Bha
APPLICANT: Padigaru, Muraidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Futeak, Katarzyna
APPLICANT: Groose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072.012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514

```

; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 255
; LENGTH: 2200
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-012-255

Query Match      95.6%; Score 1906.8; DB 13; Length 2200;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 45; Gaps 4

QY      1  ATGCCCATGAGATGATTTGGAACCTCAATTGTTACTGAGAGGTGGTGGCTCTGCTGGAA 60
DB      61  ATGGCCCATGAGATGATTTGGAACCTCAATTGTTACTGAGAGGTGGTGGCTCTGCTGGAA 120

QY      61  AGTGAACCGGAAAAAGTGTGCTTAATTGATAGCCGGCCATTGTGGAATATCAATCATCC 120
DB      121  AGTGAACCGGAAAAAGTGTGCTTAATTGATAGCCGGCCATTGTGGAATATCAATCATCC 180

QY      121  CACATTTTGGAGGCCATTATATATCACTGCTCCAAAGCTTATGAACGGAAGTTGGCAACG 180
DB      181  CACATTTTGGAGGCCATTATATATCACTGCTCCAAAGCTTATGAACGGAAGTTGGCAACG 240

QY      181  GACAAAGTGTTAATTACAGAGCTATCCAGATTACGGGAACATA-----AG 228
DB      241  GACAAAGTGTTAATTACAGAGCTATCCAGATTACGGGAACATAAGTAAACGCTCAG 300

QY      229  GTTGAACATTAATTCAGTCAGAGAGGTGTAGTTACATCAAGATCCCAAGATGTGGC 288
DB      301  GTTGAACATTAATTCAGTCAGAGAGGTGTAGTTACATCAAGATCCCAAGATGTGGC 360

QY      301  TCTCTCTCTTCAGACTGTTTCTCACTGTACTTCGGGTAAACTGAGAAAGCTTCAAC 348
DB      361  TCTCTCTCTTCAGACTGTTTCTCACTGTACTTCGGGTAAACTGAGAAAGCTTCAAC 420

QY      349  TCTGTTCACTGCTTGC-----AGTGGGTTTGTGAATTCTCTGTTGTTTC 396
DB      421  TCTGTTCACTGCTTGCAGGTTTATCTTAGTGGGTTTGTGAATTCTCTGTTGTTTC 480

QY      397  CTTGGCCTCTGTGAAGGAAAAATCCACTGTAAGTCCCTACCTGCAATTTCTCAACCTTGCCTTA 456
DB      481  CTTGGCCTCTGTGAAGGAAAAATCCACTGTAAGTCCCTACCTGCAATTTCTCAACCTTGCCTTA 540

QY      457  CCTGTTCGCAACATTTGGGCCCAACCCGAATTTTCCCAATCTTTATCTTGGCTGCGACGGA 516
DB      541  CCTGTTCGCAACATTTGGGCCCAACCCGAATTTTCCCAATCTTTATCTTGGCTGCGACGGA 600

QY      517  GATGTCTCTCAACAGAGAGCTGATACAGCAGAAATGGAAATGGTATGTGTTAAATGCCAGC 576
DB      601  GATGTCTCTCAACAGAGAGCTGATACAGCAGAAATGGAAATGGTATGTGTTAAATGCCAGC 660

QY      577  TATATCCTGTCCAAAGCCTGAATTTATCCCGAGTCTATTTCCCGCGTGGCCCTGTGAT 636
DB      661  TATATCCTGTCCAAAGCCTGAATTTATCCCGAGTCTATTTCCCGCGTGGCCCTGTGAT 720

```

637 GACAGCTTTTGAGAAATTTTGCCGTGGTTGACAATAAGTAGATTTTATT----- 690
721 GACAGCTTTTGAGAAAATTTTGCGGTGTTGGCAAAATCAGTAGATTTCAATGGTAAAG 780
691 -----GAGAAAGCAAAAGCTTCAATGAGTGTGTTCAATGCACTGTTTAGCT 738
781 TTGACTTATACAGAGAAAGCAAAAGCTTCAATGAGTGTGTTCAATGCACTGTTTAGCT 840
739 GGGATCTCCCGCTCCGCCCAACATGCTATACATGAGAGAGAGATGACATGCT 798
841 GGGATCTCCCGCTCCGCCCAACATGCTATACATGAGAGAGATGACATGCT 900
799 TTAGATGAAGCTTAC---AGATTGTGAAAGAAAAAGACTTACTATATCTCCAACTTC 855
901 TTGATGAGAGCTTACAGAGATTTGTGAAAAGAAAAAGACTTACTATATCTCCAACTTC 960
856 AATTTTCTGGGCAACTCTGCACTATGAGAAAGATTTAAGAACCAAGACTGAGACATCA 915
961 AATTTTCTGGGCAACTCTGCACTATGAGAAAGATTTAAGAACCAAGACTGAGACATCA 1020
916 GGGCCCAAGAGCAATCTCAAGCTGCTCACTGAGAGAGCAAAATGAACCTGCTCCCTGCT 975
1021 GGGCCCAAGAGCAATCTCAAGCTGCTCACTGAGAGAGCAAAATGAACCTGCTCCCTGCT 1080
976 GTCTCAGAGGGTGAACAAGAGGAGACGCCCTCTAGTCAACCTGTGCGCACTGCT 1035
1081 GTCTCAGAGGGTGAACAAGAGGAGACGCCCTCTAGTCAACCTGTGCGCACTGCT 1140
1036 AACTCAGAGGGTGAACAAGAGGAGACGCCCTCTAGTCAACCTGTGCGCACTGCT 1095
1141 AACTCAGAGGGTGAACAAGAGGAGACGCCCTCTAGTCAACCTGTGCGCACTGCT 1200
1096 GTGAGCGCGTGGCTGTTAAGAGACACCGCGGTGATCAGCGGCTCAAGTGGCTGCACTTG 1155
1201 GTGAGCGCGTGGCTGTTAAGAGACACCGCGGTGATCAGCGGCTCAAGTGGCTGCACTTG 1260
1156 TCCGAGAGAGGGTGAACAAGAGGAGACGCCCTCTAGTCAACCTGTGCGCACTGCT 1215
1261 TCCGAGAGAGGGTGAACAAGAGGAGACGCCCTCTAGTCAACCTGTGCGCACTGCT 1320
1216 TCGATTTCAATTTACGCCAGATGAGCAACATCTTCAATGAGCTTCTCTCTGAGATTCAA 1275
1321 TCGATTTCAATTTACGCCAGATGAGCAACATCTTCAATGAGCTTCTCTCTGAGATTCAA 1380
1276 GCTTTGAGATTAACAACCTTCCACTTCTGAGTGGACCAAGCTATGCGAGTTC 1335
1381 GCTTTGAGATTAACAACCTTCCACTTCTGAGTGGACCAAGCTATGCGAGTTC 1440
1336 TCCCGCTTCAAGAACTATCGAGACAGTCCGAAACCAAGTCTGATTAAGAGGAGAAAGC 1395
1441 TCCCGCTTCAAGAACTATCGAGACAGTCCGAAACCAAGTCTGATTAAGAGGAGAAAGC 1500
1396 AGCATCCCAAGAGCTGACAGACCGCGAGGCTTCAAGACGCAAGCAAGCATTTGAT 1455
1501 AGCATCCCAAGAGCTGACAGACCGCGAGGCTTCAAGACGCAAGCAAGCATTTGAT 1560
1456 TCGGTGAGAACCAACACAGATGGACACCGCCAGAGGCTCTTTTATCTCACTGATCGA 1515
1561 TCGGTGAGAACCAACACAGATGGACACCGCCAGAGGCTCTTTTATCTCACTGATCGA 1620
1516 AGTGGAGCGGTGAAGACATTTCAACAACAGCTTCTTTTCGAGCTTTTCAACAGCAG 1575
1621 AGTGGAGCGGTGAAGACATTTCAACAACAGCTTCTTTTCGAGCTTTTCAACAGCAG 1680
1576 CAGCACTCAAGAAAGCTGCGCTGGCCCTGGGCTTTAAGGCTGACCTGAGATATCTTGACC 1635
1681 CAGCACTCAAGAAAGCTGCGCTGGCCCTTTAAGGCTGACCTGAGATATCTTGACC 1740
1636 CCGCAAGCTCTAACCCCTTCCCTGAACAGAGCTGATTTTTCGACAGAGCTCTCAAC 1695
1741 CCGCAAGCTCTAACCCCTTCCCTGAACAGAGCTGATTTTTCGACAGAGCTCTCAAC 1800

```

QY      1696 TTCTACTTGCCTCAGCCATCTAAGGAGGCGAGGCCAGTACTCTGCTTACAGCTGCAAGC 1755
Db      1801 TTCTACTTGCCTCAGCCATCTAAGGAGGCGAGGCCAGTACTCTGCTTACAGCTGCAAGC 1860
QY      1756 CACTGCGCACTTGGGGAGACCAAGCTTATTTCTGTGCGCAGCGCGCAGAAAGCCAAAGTAC 1815
Db      1861 CAGCTGCCACTTGGGGAGACCAAGCTTATTTCTGTGCGCAGCGCGCAGAAAGTAC 1920
QY      1816 AAGCTGACTGCGCGCGAGCTGSCATGAAAGAGAGCCCTTTGAAAGACAGTTTAAAGCG 1875
Db      1921 AAGGCTGACTGCGCGCGAGCTGSCATGAAAGAGAGCCCTTTGAAAGACAGTTTAAAGCG 1980
QY      1876 AAGAGCTGCCAAAGGAATTTGAGAGAGCATCATGTCAAGAAACAGTCCAGGAAAGAG 1935
Db      1981 AAGAGCTGCCAAAGGAATTTGAGAGAGCATCATGTCAAGAAACAGTCCAGGAAAGAG 2040
QY      1936 CTGGGAAAGTGGGCACTAGTACTTTTGGGCAACATGGAATCATTTGAGTCTCC 1995
Db      2041 CTGGGAAAGTGGGCACTAGTACTTTTGGGCAACATGGAATCATTTGAGTCTCC 2100

RESULT 15
US-09-964-277-20
; Sequence 20, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964.277
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 3332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-20

Query Match      82.9%; Score 1653.8; DB 9; Length 3332;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1829; Conservative 0; Mismatches 2; Indels 164; Gaps 1;

QY      1 AAGGCCCACTGAGATGATTTGAATCTCAATTTGTACTGAGAGGTTGTGGCTCTGGTGGAA 60
Db      562 AATGGCCCACTGAGATGATTTGAATCTCAATTTGTACTGAGAGGTTGTGGCTCTGGTGGAA 621
QY      61 AATGGAACGGAAAAAGTGTGCTGAATTTGATGCGCGCATTTGTGGAAATCAATTCATCC 120
Db      622 AATGGAACGGAAAAAGTGTGCTGAATTTGATGCGCGCATTTGTGGAAATCAATTCATCC 681
QY      121 CACATTTTGAAGCCCTTAATATATCAACTGCTTCAAGCTTATGAAAGCAAGTGTGCAAGC 180
Db      682 CACATTTTGAAGCCCTTAATATATCAACTGCTTCAAGCTTATGAAAGCAAGTGTGCAAGC 741
QY      181 GACAAAGTGTAAATTACAGAGCTATCCAGATTCACGGAACATTAAGTTGACATTGAT 240
Db      742 GACAAAGTGTAAATTACAGAGCTATCCAGATTCACGGAACATTAAGTTGACATTGAT 801
QY      241 TGCAGTCAAGAGGTTGATAGTTTAAGATCAAGCTCCCAAGATTTGCTGCTCTCTTCA 300
Db      802 TGCAGTCAAGAGGTTGATAGTTTAAGATCAAGCTCCCAAGATTTGCTGCTCTCTTCA 861
QY      301 GACTGTTTCTACCTGTACTCTGGGTAAACTGGAGAAAGCTTCAACTGTGTCACCTG 360
Db      862 GACTGTTTCTACCTGTACTCTGGGTAAACTGGAGAAAGCTTCAACTGTGTCACCTG 921
QY      361 CTTCGAGGTGGGTTGTCTGAGTTCTCTCGTTGTTTCTGAGCTCTGTGAAGAAATCC 420
Db      922 CTTCG----- 926
QY      421 ACTCTAGTCCCTACCTGATTTTCTCAGCCTTGCTTAACTGTGCAACATTTGGGCAACC 480

```

Db	927	-----	926	Db	1898	TCTCCATGATCGAAGTGGAGCGTGGAGAGCAATTAACCAACAGCTTCTTTGGC	1957	
Qy	481	CGAATCTTCCCAATCTTATCTTGGCTGCAGAGAGATGCTCTCAACAAAGAGCTGATA	540	Qy	1561	CTTTCCACAGCCAGAGACCTCAAGAACTCTGCTGGGCTTGAGGCTGGCAC	1620	
Db	927	-----	AGGAGCTGATG	937	Db	1958	CTTTCCACAGCCAGAGACCTCAAGAACTCTGCTGGGCTTGAGGCTGGCAC	2017
Qy	541	CAGCAGAAATGGGATGGTATGTTAAATGCAAGCTATACTGTCCAAAGCCGTGCTT	600	Qy	1621	TCGGATATCTTGGAGCCCGCCAGACCTTAACCTTCCCTGACACAGCAGTGTATTTGGC	1680	
Db	938	CACACAGAAATGGGATGGTATGTTAAATGCAAGCTATACTGTCCAAAGCCGTGCTT	997	Db	2018	TCGGATATCTTGGAGCCCGCCAGACCTTAACCTTCCCTGACACAGCAGTGTATTTGGC	2077	
Qy	601	ATCCCGAGTCTGATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG	660	Qy	1681	ACAGATCTCTCACTTCTACTCTGCTCAGCCATCTACGAGGAGAGTCCAGTTACTT	1740	
Db	998	ATCCCGAGTCTGATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG	1057	Db	2078	ACAGATCTCTCACTTCTACTCTGCTCAGCCATCTACGAGGAGAGTCCAGTTACTT	2137	
Qy	661	CCGTGGTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTTCATGATGTT	720	Qy	1741	GCCTAAGCTGACAGCCAGCTGCCCACTTGGGAGAACAGTCTATCTGTGGCAGGCGG	1800	
Db	1058	CCGTGGTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTTCATGATGTT	1117	Db	2138	GCCTAAGCTGACAGCCAGCTGCCCACTTGGGAGAACAGTCTATCTGTGGCAGGCGG	2197	
Qy	721	CTAGTGACCTTTAGCTGGGATCTCCCGTCCGCGCACTGCTATGCGCTACATCATG	780	Qy	1801	CAGAACCAAGTGAAGAGCTGACTGCGCGGAGCTGGCATGAAAGAGACCCCTTGAA	1860	
Db	1118	CTAGTGACCTTTAGCTGGGATCTCCCGTCCGCGCACTGCTATGCGCTACATCATG	1177	Db	2198	CAGAACCAAGTGAAGAGCTGACTGCGCGGAGCTGGCATGAAAGAGACCCCTTGAA	2257	
Qy	781	AAGAGATGAGCATGTCTTATGATGAAGCTTACAGATTTGTGAAAAGAAAGACCTACT	840	Qy	1861	AAGCAGTTTAAACGAGAAAGCTGCCAATGAAATTTGAGAGAGATCATGTGAGAAC	1920	
Db	1178	AAGAGATGAGCATGTCTTATGATGAAGCTTACAGATTTGTGAAAAGAAAGACCTACT	1237	Db	2258	AAGCAGTTTAAACGAGAAAGCTGCCAATGAAATTTGAGAGAGATCATGTGAGAAC	2317	
Qy	841	ATACTTCCAACTTCAATTTTCTGGGCCAACTCCTGACTATGAGAAAGATTAAAGAC	900	Qy	1921	AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTAGCTTTTGGGAGAGATGAA	1980	
Db	1238	ATACTTCCAACTTCAATTTTCTGGGCCAACTCCTGACTATGAGAAAGATTAAAGAC	1297	Db	2318	AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTAGCTTTTGGGAGAGATGAA	2377	
Qy	901	CAGACTGGAGCATCAAGGCGCAAAAGAGCAACTCAAGTGTGCACTGTGGAAGCCAAAT	960	Qy	1981	ATCATTAGAGTCTCC	1995	
Db	1298	CAGACTGGAGCATCAAGGCGCAAAAGAGCAACTCAAGTGTGCACTGTGGAAGCCAAAT	1357	Db	2378	ATCATTAGAGTCTCC	2392	
Qy	961	GAACCTGTCCCTGCTGTCTGAGAGGTGACAGAAAAGCGAGCGCCCTCAGTCCACC	1020					
Db	1358	GAACCTGTCCCTGCTGTCTGAGAGGTGACAGAAAAGCGAGCGCCCTCAGTCCACC	1417					
Qy	1021	TGTGCCGACTCTGTACTACCTTCAAGAGCAGAGCAAAAGCCCGTGCATCCCGCAGCGTG	1080					
Db	1418	TGTGCCGACTCTGTACTACCTTCAAGAGCAGAGCAAAAGCCCGTGCATCCCGCAGCGTG	1477					
Qy	1081	CCCAAGCTGCCAGCGTGCAGCGCTGCTGTAGAGACAGCCCGCTGTGTAACAGCGCTC	1140					
Db	1478	CCCAAGCTGCCAGCGTGCAGCGCTGCTGTAGAGACAGCCCGCTGTGTAACAGCGCTC	1537					
Qy	1141	AGTGGCTGACCGTCCGCGAGACAGGCTGAAAGACAGCAATAGCTCAAGCGTTCCTTC	1200					
Db	1538	AGTGGCTGACCGTCCGCGAGACAGGCTGAAAGACAGCAATAGCTCAAGCGTTCCTTC	1597					
Qy	1201	TCTCTGGATATCAAAATCAGTTTCAATTAAGCCAGCATGGCAGCATCTTAATGAGCTTC	1260					
Db	1598	TCTCTGGATATCAAAATCAGTTTCAATTAAGCCAGCATGGCAGCATCTTAATGAGCTTC	1657					
Qy	1261	TCTCTCATCAGAAAGTGTGTAATTAATAAACTTCCACTACTGTGATGGAGCCAAC	1320					
Db	1658	TCTCTCATCAGAAAGTGTGTAATTAATAAACTTCCACTACTGTGATGGAGCCAAC	1717					
Qy	1321	AAGCTATGCGAGTTCTCCCGTTCAGAACTATCGAGAGAGACTCCCGAAAACAGTCTC	1380					
Db	1718	AAGCTATGCGAGTTCTCCCGTTCAGAACTATCGAGAGAGACTCCCGAAAACAGTCTC	1777					
Qy	1381	GATTAAGAGAGAGCAGACATCCCAAGAGCTGACAGCCGCAAGGCTTTCAGACAGCCAG	1440					
Db	1778	GATTAAGAGAGAGCAGACATCCCAAGAGCTGACAGCCGCAAGGCTTTCAGACAGCCAG	1837					
Qy	1441	AGCAAGGATTTGCAATTTGGGTCAAGAACAGAGAGAGTGGCAACCGCCAGAGAGTCCCTTTTA	1500					
Db	1838	AGCAAGGATTTGCAATTTGGGTCAAGAACAGAGAGAGTGGCAACCGCCAGAGAGTCCCTTTTA	1897					
Qy	1501	TCTCACTGATCGAAGTGGAGAGCTGAGAGACAAATTAACACAGCTTCTTTTGGC	1560					

Search completed: June 22, 2004, 03:25:18
Job time : 1198.09 secs

~~... This Page Blank (aspj0)~~

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 17:08:00 ; Search time 165.047 Seconds
(Without alignments)
6707.969 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccatgagatgatctg.....tggaaatcatgaggtctcc 1995

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCUTS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1991.8	99.8	1998 4	US-09-816-494-3 Sequence 3, Appl1
2	1991.8	99.8	3544 4	US-09-816-494-1 Sequence 1, Appl1
3	387.8	19.4	2377 4	US-09-920-668-3 Sequence 3, Appl1
4	223	11.2	279 4	US-09-016-434-91 Sequence 91, Appl1
5	124.6	6.2	1830 4	US-09-557-921-1 Sequence 1, Appl1
6	106.8	5.4	2303 4	US-09-922-146-3 Sequence 3, Appl1
7	96.2	4.8	2109 4	US-09-016-434-1135 Sequence 1135, Ap
8	96.2	4.8	2109 4	US-09-023-655-946 Sequence 946, App
9	94.6	4.7	1208 4	US-09-023-655-347 Sequence 347, App
10	93.6	4.7	1619 4	US-09-702-705-801 Sequence 801, App
11	93.6	4.7	1619 4	US-09-736-457-801 Sequence 801, App
12	93.6	4.7	1619 4	US-09-614-1248-801 Sequence 801, App
13	93.6	4.7	1619 4	US-09-614-1248-801 Sequence 801, App
14	93.6	4.7	1619 4	US-09-589-184-801 Sequence 801, App
15	93.6	4.7	4637 4	US-09-702-705-804 Sequence 804, App
16	93.6	4.7	4637 4	US-09-736-457-804 Sequence 804, App
17	93.6	4.7	4637 4	US-09-614-1248-804 Sequence 804, App
18	93.6	4.7	4637 4	US-09-614-1248-804 Sequence 804, App
19	93.6	4.7	4637 4	US-09-589-184-804 Sequence 804, App
20	92	4.6	1238 2	US-08-530-290-11 Sequence 11, Appl
21	92	4.6	1238 4	US-09-702-705-803 Sequence 803, App
22	92	4.6	1238 4	US-09-736-457-803 Sequence 803, App
23	92	4.6	1238 4	US-09-614-1248-803 Sequence 803, App
24	92	4.6	1238 4	US-09-671-325-803 Sequence 803, App
25	92	4.6	1238 4	US-09-589-184-803 Sequence 803, App
26	92	4.6	2064 4	US-09-702-705-825 Sequence 825, App
27	92	4.6	2064 4	US-09-736-457-825 Sequence 825, App

28	92	4.6	2064 4	US-09-614-1248-825 Sequence 825, App
29	92	4.6	2064 4	US-09-671-325-825 Sequence 825, App
30	92	4.6	2064 4	US-09-589-184-825 Sequence 825, App
31	92	4.6	2109 4	US-09-702-705-826 Sequence 826, App
32	92	4.6	2109 4	US-09-736-457-826 Sequence 826, App
33	92	4.6	2109 4	US-09-614-1248-826 Sequence 826, App
34	92	4.6	2109 4	US-09-671-325-826 Sequence 826, App
35	92	4.6	2109 4	US-09-589-184-826 Sequence 826, App
36	92	4.6	2240 4	US-09-016-434-1100 Sequence 1100, Ap
37	89.4	4.5	2000 4	US-09-016-434-1291 Sequence 1291, App
38	86.2	4.3	240 4	US-09-016-434-776 Sequence 776, App
39	85.6	4.3	1987 2	US-08-990-379-1 Sequence 1, Appl1
40	85.6	4.3	1993 2	US-08-990-379-2 Sequence 2, Appl1
41	80.4	4.0	944 4	US-09-371-671B-10 Sequence 10, Appl
42	79	4.0	539 4	US-09-389-681-311 Sequence 311, App
43	79	4.0	539 4	US-09-620-405B-311 Sequence 311, App
44	79	4.0	539 4	US-09-339-338-311 Sequence 311, App
45	79	4.0	539 4	US-09-433-826B-311 Sequence 311, App

ALIGNMENTS

```
RESULT 1
US-09-816-494-3
; Sequence 3, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-816-494-3

Query Match      99.8%; Score 1991.8; DB 4; Length 1998;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ATGGCCCATGAGATGATGGAATCTCAAAATTGTAAGAGAGTTGGGCTGCTGGAA 60
      1  ATGGCCCATGAGATGATGGAATCTCAAAATTGTAAGAGAGTTGGGCTGCTGGAA 60
DB
QY      61  AGTGAACGGAAGAAAGTGTCTGCTAATTTGATAGCCGCGCATTTTGGAAATCAATATCC 120
      61  AGTGAACGGAAGAAAGTGTCTGCTAATTTGATAGCCGCGCATTTTGGAAATCAATATCC 120
DB
QY      121  CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATGAAGAGAGTTGCAACAG 180
      121  CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATGAAGAGAGTTGCAACAG 180
DB
QY      181  GACAAGGTGTAATTCAGAGCTCAGCATTCAGCGAAATTAAGAGTTGACATTTGAT 240
      181  GACAAGGTGTAATTCAGAGCTCAGCATTCAGCGAAATTAAGAGTTGACATTTGAT 240
DB
QY      241  TGCAGTCAGAGAGTTGATTTACGATCAAGCTCCCAAGATGTCCTCTCTCTCA 300
      241  TGCAGTCAGAGAGTTGATTTACGATCAAGCTCCCAAGATGTCCTCTCTCTCA 300
DB
QY      301  GACTGTTTCTCAGTCTGTAATTCTGGGTAACAGAGAGCTTCAACTGTTCACCTG 360
      301  GACTGTTTCTCAGTCTGTAATTCTGGGTAACAGAGAGCTTCAACTGTTCACCTG 360
DB
QY      361  CTGCAAGTGGGTTTCTGCTGATGTTCTCTGTTTCTCTGCTCTGTAAGAAATCC 420
      361  CTGCAAGTGGGTTTCTGCTGATGTTCTCTGTTTCTCTGCTCTGTAAGAAATCC 420
```

Db 361 CTGAGAGTGGGTTTGAGAGTCTCTGTTGTTCCCTGGCCCTGTGAGAGAAATCC 420
Qy 421 ACTCTAGTCCCTTACCTGATTTCTCAGCCTTGCTACCTGTGCAACATTTGGGCCAAC 480
Db 421 ACTCTAGTCCCTTACCTGATTTCTCAGCCTTGCTACCTGTGCAACATTTGGGCCAAC 480
Qy 481 CGAATTTCTCCCATCTTTATCTTGGTCCAGCGAGATGTCTCTCAACAGAGCTGATG 540
Db 481 CGAATTTCTCCCATCTTTATCTTGGTCCAGCGAGATGTCTCTCAACAGAGCTGATG 540
Qy 541 CAGAGAAATGGATTGGTTATGTGTTAAATGCAAGCTATAGCTGTCCAAACCTGACTT 600
Db 541 CAGAGAAATGGATTGGTTATGTGTTAAATGCAAGCTATAGCTGTCCAAACCTGACTT 600
Qy 601 ATCCCGAGTCTCATTTCCGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Db 601 ATCCCGAGTCTCATTTCCGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Qy 661 CCGTGTGGAGCAATCAGTAGATTTCATGAGAAAGCAAAAGCTCCATATGATGTTGT 720
Db 661 CCGTGTGGAGCAATCAGTAGATTTCATGAGAAAGCAAAAGCTCCATATGATGTTGT 720
Qy 721 CTAGTGACTGTTTGTGCTGGGATCTCCCGCTCCGCAACATGCTATGCTATCATG 780
Db 721 CTAGTGACTGTTTGTGCTGGGATCTCCCGCTCCGCAACATGCTATGCTATCATG 780
Qy 781 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAACTTACT 840
Db 781 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAACTTACT 840
Qy 841 ATATCTCCAACTTCAATTTTCTGGGGCAACTCTGGAATATGAGAAAGATTAAGAAC 900
Db 841 ATATCTCCAACTTCAATTTTCTGGGGCAACTCTGGAATATGAGAAAGATTAAGAAC 900
Qy 901 CAGACTGGAGATCAGGGGCCAAAGAGCAATCTCAAGCTGCTGCACTGTGAGAGCAAT 960
Db 901 CAGACTGGAGATCAGGGGCCAAAGAGCAATCTCAAGCTGCTGCACTGTGAGAGCAAT 960
Qy 961 GAACCTGTCTCTGCTCTCAGAGGGTGAAGAAAGCGAGAGCCCTTGACATGCC 1020
Db 961 GAACCTGTCTCTGCTCTCAGAGGGTGAAGAAAGCGAGAGCCCTTGACATGCC 1020
Qy 1021 TGTGCGCACTGTCTCTCAGAGGGTGAAGAAAGCGAGAGCCCTTGACATGCC 1080
Db 1021 TGTGCGCACTGTCTCTCAGAGGGTGAAGAAAGCGAGAGCCCTTGACATGCC 1080
Qy 1081 CCCAGCGTCCAGAGGTGAGCGTCTGTTAAGAGACAGCCCGCTGTACAGCGCTC 1140
Db 1081 CCCAGCGTCCAGAGGTGAGCGTCTGTTAAGAGACAGCCCGCTGTACAGCGCTC 1140
Qy 1141 AGTGGAGTGCACCTGTCCGAGACAGGCTGGAACAAGCAATAGCTCAAGCGTCTC 1200
Db 1141 AGTGGAGTGCACCTGTCCGAGACAGGCTGGAACAAGCAATAGCTCAAGCGTCTC 1200
Qy 1201 TCTCTGATATCAATCAGTTTCAATTTACGCGAGCTGAGAGATCTTAACTGCTTC 1260
Db 1201 TCTCTGATATCAATCAGTTTCAATTTACGCGAGCTGAGAGATCTTAACTGCTTC 1260
Qy 1261 TCTCTGATATCAATCAGTTTCAATTTACGCGAGCTGAGAGATCTTAACTGCTTC 1320
Db 1261 TCTCTGATATCAATCAGTTTCAATTTACGCGAGCTGAGAGATCTTAACTGCTTC 1320
Qy 1321 AAGCTATGCAAGTCTCCCTGTTCAAGAACTATGAGACAGTCTCCGAAACAGTCT 1380
Db 1321 AAGCTATGCAAGTCTCCCTGTTCAAGAACTATGAGACAGTCTCCGAAACAGTCT 1380
Qy 1381 GATPAGAGGAAAGCCGACATCCCAAGAGCTGACAGCCGAGCTTTCAGAGCGAG 1440
Db 1381 GATPAGAGGAAAGCCGACATCCCAAGAGCTGACAGCCGAGCTTTCAGAGCGAG 1440
Qy 1441 AGCAACGATGTCATTCGCTCAGAAACAGCGAGCTGACAGCGCTTCCCTTTT 1500
Db 1441 AGCAACGATGTCATTCGCTCAGAAACAGCGAGCTGACAGCGCTTCCCTTTT 1500

Db 1441 AGCAACGATGTCATTCGCTCAGAAACAGCGAGCTGACAGCGCTTCCCTTTT 1500
Qy 1501 TCTCCACTGATTCGAAATGAGAGCGTGTGAGACAATTAACACAGAGCTTCTTTGGC 1560
Db 1501 TCTCCACTGATTCGAAATGAGAGCGTGTGAGACAATTAACACAGAGCTTCTTTGGC 1560
Qy 1561 CTTTCACAGCAGCAGACACACCTTCAAGAGCTGCTGCTGCTTAAAGGCTGGCAC 1620
Db 1561 CTTTCACAGCAGCAGACACACCTTCAAGAGCTGCTGCTGCTTAAAGGCTGGCAC 1620
Qy 1621 TGGATATCTGGGCCCCCAGACCTTACCTTCCCTGACAGAGCTGATTTTGGC 1680
Db 1621 TGGATATCTGGGCCCCCAGACCTTACCTTCCCTGACAGAGCTGATTTTGGC 1680
Qy 1681 ACAGAGTCTCAGACTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db 1681 ACAGAGTCTCAGACTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Qy 1741 GCTTACAGCTGACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1741 GCTTACAGCTGACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Qy 1801 CAGAGCAAGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1801 CAGAGCAAGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Qy 1861 AAGAGTTTAAAGCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1861 AAGAGTTTAAAGCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Qy 1921 AGGTCAAGGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 1921 AGGTCAAGGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Qy 1981 ATCATTTAGGCTGCTC 1995
Db 1981 ATCATTTAGGCTGCTC 1995

RESULT 2
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589)...(2583)
US-09-816-494-1

Query Match 99.8%; Score 1991.8; DB 4; Length 3544;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGCGCCATGAGATGATTTGAACTCAATTTGTAATGAGAGGTTGGCTGTGAGAA 60
Db 589 ATGCGCCATGAGATGATTTGAACTCAATTTGTAATGAGAGGTTGGCTGTGAGAA 648
Qy 61 AGTGAACGAAAAAGTGTCTAATTTGATAGCGGCGCTTTGTGAATCAATCATCC 120
Db 61 AGTGAACGAAAAAGTGTCTAATTTGATAGCGGCGCTTTGTGAATCAATCATCC 120

Dh 649 AGTGAACGAAAAAGTCTGCTAATTGATAGCCGGCATTTGNGAATAACAATCATCC 708
Qy 121 CACATTTTGGAAAGCATTAAATATCACTGCTCCAGTTATYGAAGGAGTTGCAACAG 180
Db 709 CACATTTTGGAAAGCATTAAATATCACTGCTCCAGTTATYGAAGGAGTTGCAACAG 768
Qy 181 GACAAAGTGTAAATTCAGAGCTCATCAGACATTCAGCGAAACAATAGATTGACATTGAT 240
Db 769 GACAAAGTGTAAATTCAGAGCTCATCAGACATTCAGCGAAACAATAGATTGAT 828
Qy 241 TGCAGTCAGAAAGTGTAGTTAGATCAAAAGCTCCAGAGATGTGCTCTCTTCA 300
Db 829 TGCAGTCAGAAAGTGTAGTTAGATCAAAAGCTCCAGAGATGTGCTCTCTTCA 888
Qy 301 GACTGTTTTTCTCACTGTAATTTGGGTAACCTGGAGAAAGCTTCAACTCTTCACTG 360
Db 889 GACTGTTTTTCTCACTGTAATTTGGGTAACCTGGAGAAAGCTTCAACTCTTCACTG 948
Qy 361 CTGTCAGAGTGGGTTTGCTGAGTTCTCTGCTGTTTCCCTGGCCCTGTGAGAAATCC 420
Db 949 CTGTCAGAGTGGGTTTGCTGAGTTCTCTGCTGTTTCCCTGGCCCTGTGAGAAATCC 1008
Qy 421 ACTTAGTCCCTCACTGCACTTCTCAGCCCTTGCTTACCTGTTGCAACATTTGGGCCAAC 480
Db 1009 ACTTAGTCCCTCACTGCACTTCTCAGCCCTTGCTTACCTGTTGCAACATTTGGGCCAAC 1068
Qy 481 CGAATTTCTTCCAACTTTATCTTGGCTGCTCAGCGAGATGTCTCAACAAGAGCTGATA 540
Db 1069 CGAATTTCTTCCAACTTTATCTTGGCTGCTCAGCGAGATGTCTCAACAAGAGCTGATA 1128
Qy 541 CACACAGAAATGGGATGTTATGTTTAAATGACAGCATATCCTGCCAAAGCCCTGACTTT 600
Db 1129 CACACAGAAATGGGATGTTATGTTTAAATGACAGCATATCCTGCCAAAGCCCTGACTTT 1188
Qy 601 ATCCCGAGTCTCATTTCTGCGTGCTGCTGATGATGACAGCTTTGTTGAGAAATTTTG 660
Db 1189 ATCCCGAGTCTCATTTCTGCGTGCTGCTGATGATGACAGCTTTGTTGAGAAATTTTG 1248
Qy 661 CCGTGTGGGACAAATCATGATATTTTATGAGAAAGCAAAAGCCTTCAATGATGTT 720
Db 1249 CCGTGTGGGACAAATCATGATATTTTATGAGAAAGCAAAAGCCTTCAATGATGTT 1308
Qy 721 CTGATGACCTTTTACGTGGGATCTCCGCTCCGCCCACTGCTATGCTATCATATG 780
Db 1309 CTGATGACCTTTTACGTGGGATCTCCGCTCCGCCCACTGCTATGCTATCATATG 1368
Qy 781 AAGAGATGACATGTCTTAGATGAGCTTACAGATTTTGTGAAGAAAGAACTACT 840
Db 1369 AAGAGATGACATGTCTTAGATGAGCTTACAGATTTTGTGAAGAAAGAAAGCTACT 1428
Qy 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
Db 1429 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAGAAC 1488
Qy 901 CAGACTGAGACATCAAGGCCCAAGAGCAAACTCAAGCTGTGACCTGAGAAAGCCCAAT 960
Db 1489 CAGACTGAGACATCAAGGCCCAAGAGCAAACTCAAGCTGTGACCTGAGAAAGCCCAAT 1548
Qy 961 GAACCTGTCTGCTGCTCAGAGAGTGAAGAGAAAGCAGAGAGCCCTCAGTCCACC 1020
Db 1549 GAACCTGTCTGCTGCTCAGAGAGTGAAGAGAAAGCAGAGAGCCCTCAGTCCACC 1080
Qy 1021 TGTGCGCACTCTGCTACTCAGAGAGCAAGCAAAAGCCCTGATCCCGCAGGCTG 1080
Db 1609 TGTGCGCACTCTGCTACTCAGAGAGCAAGCAAAAGCCCTGATCCCGCAGGCTG 1668
Qy 1081 CCGAGCGTGGCCAGCGGTCAGCGCTGCTGTTAGAGACAGCCCGCTGTGACAGCGCTC 1140
Db 1669 CCGAGCGTGGCCAGCGGTCAGCGCTGCTGTTAGAGACAGCCCGCTGTGACAGCGCTC 1728
Qy 1141 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGCTTCTTTC 1200
Db 1729 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGCTTCTTTC 1788

Qy 1201 TCTCTGGATATGAATCAGTTTCAATATTCAGGCGAGATGGAGCATCTTACATGGCTTC 1260
Db 1789 TCTCTGGATATGAATCAGTTTCAATATTCAGGCGAGATGGAGCATCTTACATGGCTTC 1848
Qy 1261 TCTCTCATCAGAAAGATCTTGGAAATCTAACAACCTTCCACTGATCTGTGAATGGAAAC 1320
Db 1849 TCTCTCATCAGAAAGATCTTGGAAATCTAACAACCTTCCACTGATCTGTGAATGGAAAC 1908
Qy 1321 AAGCTATGCCAGTTCTTCCCTGTTCAAGAACTATGAGAGACATCTCCGAAACCATGCT 1380
Db 1909 AAGCTATGCCAGTTCTTCCCTGTTCAAGAACTATGAGAGACATCTCCGAAACCATGCT 1968
Qy 1381 GATTAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGACACCGCCAGGCTTACAGACGAC 1440
Db 1969 GATTAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGACACCGCCAGGCTTACAGACGAC 2028
Qy 1441 AGCAAGCATGTCATTCGGTCAAGAACCAAGCATGAGACCGCCAGAGATCTCTTTTA 1500
Db 2029 AGCAAGCATGTCATTCGGTCAAGAACCAAGCATGAGACCGCCAGAGATCTCTTTTA 2088
Qy 1501 TCTCAGCTGATCAGAAATGAGAGCTGAGAGACATTAACAACAGCTTCTTTTGGC 1560
Db 2089 TCTCAGCTGATCAGAAATGAGAGCTGAGAGACATTAACAACAGCTTCTTTTGGC 2148
Qy 1561 CTTTCCAGCAGCAGACAGACCTCAGAAAGTCTGCTGGGCTGGAAGGCTGGCAC 1620
Db 2149 CTTTCCAGCAGCAGACAGACCTCAGAAAGTCTGCTGGGCTGGAAGGCTGGCAC 2208
Qy 1621 TCGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACCAAGCTGATTTTGGC 1680
Db 2209 TCGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACCAAGCTGATTTTGGC 2268
Qy 1681 ACAGATCTCTCAGACTTCTACTCTGCTTACCCCTTACCCCTTACCAAGCTGATTTTGGC 1740
Db 2269 ACAGATCTCTCAGACTTCTACTCTGCTTACCCCTTACCCCTTACCAAGCTGATTTTGGC 2328
Qy 1741 GCTTACAGCTGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Db 2329 GCTTACAGCTGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388
Qy 1801 CAGAAACCAAGTACAGAGCTGATCGCGCGAGAGCTGATGAGAGAGAGAGAGAGAGAGAG 1860
Db 2389 CAGAAACCAAGTACAGAGCTGATCGCGCGAGAGAGCTGATGAGAGAGAGAGAGAGAGAG 2448
Qy 1861 AAGCAGTTTAAACGAGAAAGCTGCCAAATGAAATTTGAGAGAGCATCATGTGAGAGAAC 1920
Db 2449 AAGCAGTTTAAACGAGAAAGCTGCCAAATGAAATTTGAGAGAGCATCATGTGAGAGAAC 2508
Qy 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTACCTTTCGGGAGACATGGA 1980
Db 2509 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTACCTTTCGGGAGACATGGA 2568
Qy 1981 ATCATTGAGGTCTCC 1995
Db 2569 ATCATTGAGGTCTCC 2583

RESULT 3
US-09-920-668-3
; Sequence 3, Application US/09920668
; Patent No. 6482644
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; APPLICANT: Brett P. Morla
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 8 EXPRESSION
; FILE REFERENCE: RTS-0246
; CURRENT APPLICATION NUMBER: US/09/920,668
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 2377
; TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (135) ... (2012)
US-09-920-668-3

Query Match

Best Local Similarity 19.4%; Score 387.8; DB 4; Length 2377;
Matches 716; Conservative 0; Mismatches 452; Indels 27; Gaps 3;

```
28 ATTGTTACTGAGAGGTTGGTGGCTTCTGTGAAAAGTGAACGGAAGAAAGTGTGCTAATT 87
165 ATGATGACCAAGAACTGCGCAGCGCTGTGGGGGCGGCGCTGGGGGGCGCGTGGTATC 224
88 GATACCGGCGCATTTGTGGAATACAAATACCAATTTTGAAGCCATTAAATCAAC 147
225 GAGACCGCGCTCTTGTGTGAAGTACAGAGTGGGAGTGTCTAGCTCCGTCAACATCTGC 284
148 TCTTCAAGCTTATGAAAGCGAGTTGCAACAGAAAGTGTAAATTAACAAGCTCATC 207
285 TGCTTCAAGCTGTGTGAAGCGCGGCTGTGCAAGGCGAAGTGAACCATTTGGAGTCAATC 344
208 CAGCATTCAGCGAAACCTAAAGTTGACATTTGATGAGTCAAGAGTTGTACTTAAATGAT 267
345 CAGCGGCTGTGACGACGCGAGTGTGAGGCTAGGAGCCACAGAGAGTGTGTATGAC 404
268 CAAAGCTCCCAAGATGTGCTCTCTCTTCTTCAAGCTGTGTCTCACTGTACTTCTGGGT 327
405 CAGAGCAGCGGAGCCAGCCAGCGCTGTGGCGGAGACAGCTTCTCTCACTCTGTGAGC 464
328 AAACGTGAGAAAGAGCTTCAACTGTGTACCTGTGAGGTTGGTGTGAGTTCTCT 387
465 AAGCTGAGAGGCTGTCTTCAACAGCGTGGCATCTCACTGGGGGCTTGGCACCTTCTCC 524
388 CGTTGTTTCCCTGGGCTCTGTGAAGGAAATCCACT---CTAGTCCCTACCTGCACTTCT 444
525 TCTCTCTTCCCGGCTCTGTGAGGCGAGCGTGTGCTGCTGTACCTGAGCCCTCTCC 584
445 CAGCGTGTCTTACCTGTGCGCAACATTTGGGCGAACCAGATTTCTTCAATCTTATCTT 504
585 CAGCGCTCTGCTGTGCGCGAGGTGGGCTGACCGGATCTGTCTCACTCACTTACCTG 644
505 GCGTGTGAGAGAGTGTCTCAACAGAGAGCTATACAGCAAGTGGATGGTATGTG 564
645 GGTCTGTGAGAGAGCGTCTTAAACAGAGTGTATGACGCAAAAGGAAATGAGCTAGCTC 704
565 TTAATGCGAGCTTATCTGTCCAAAGCTGACTTATCCCGAGTCTCATTTCTGCT 624
705 CTGACGCGCAAGCACTCTGTGCGCAAGCTGACTTATCTGTGAGAGCGCTTATGCTG 764
625 GTGCTGTGATGACAGCTTTTGTGAAGAAATTTTCCGTGTGGAACAATCACTAGAT 684
765 GTCCCATCAAGCACTACTGTGAAAATCTGCTGCGCTGTGGAACAAGTCAATGAG 824
685 TTCAATGAGAAAGAAAGCCCTTCAATGATGTGTCTAGTGCAGCTTTAGCGGGATC 744
825 TTCAATGAGAAAGCAAGCTCTTCCAGCTGCCAATGATCTGCTGTGAGGCAATC 884
745 TCCCGTGTGCGCAATGCTATGCGCTCAATCAATGAGAGATGACATGTCTTAAAT 804
885 TCCCGTGTGCGCAATGCTATGCGCTCAATCAATGAGAGCAATGTCTTAAAT 864
805 GAAAGCTTCAAGTTTGTGAAGAAAGAAAGCACTTAAATCTTCAACTTCAATTTTCTG 924
945 GAGCGCTTCAAGTTTGTGAAGAGAGAGCGCGCTTCAATCTGCGCAACTTCAACTTCTG 1004
1005 GCGCAGCTGTGAGTGTGAGAGAGAGCGCGCTTCAATCTGCGCAACTTCAACTTCTG 1064
925 AGCAAACTCAAGCTGTGAGCTGTGAGAGAGCAAAATGAACTGTCTCTGTCTGAGAG 994
1065 GGCA-----CCCCCTCAAGAGAGCGCGGAGCTCTGCGCCAGTCTCTGCGCGG 1112
```

```
QY 985 GGTGACAGAAAGCCAGAGCGCCCTCACTGATCCACCTGTGCGGACTGTGCTAAGAG 1044
DB 1113 GCGCCGCTGTGCAAGGCTGTGCAACCACTTACTCAGAGAGCGCTGTCCAGAGGATGCGGCT 1172
QY 1045 GACAGAGAGCAAGAGCGCGCTGTGATCCGCGCAGCGGTGCCAGCGTCCAGCTGCAAGCCG 1104
DB 1173 GC-----CAGGAGAGCGCGCTGTGAGCGCGGAGAGCGCCCGCGCGCCCG 1220
QY 1105 TCGCTGTGAGAGAGAGCGCGCTGTGACAGCGCTCAAGTGTGCACTGTGCGGAGAGC 1164
DB 1221 AGCGCCCGGCGCAGCGGCACTGACAGAGGCGCTGTGCGGCGCTGTGACCTCTCTGAGAC 1280
QY 1165 AGCTGTGAGAGAGAGCAATTAAGCTCAAGCTTCTCTCTGTGATATCAATGAG 1219
DB 1281 CGCTGTGAGAGAGCACTTAAGCGCTTCAAGCGCTCTCTCTCTGTGATATCAATGAG 1335
```

RESULT 4

US-09-016-434-91

Sequence 91, Application US/09016434

Patent No. 650938

GENERAL INFORMATION:

APPLICANT: Janice Au-Young

APPLICANT: Jeffrey J. Selhammer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

NUMBER OF INVENTION: PATHWAY GENE EXPRESSION

CORRESPONDENCE ADDRESS:

ADDRESS: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION NUMBER: US/09/016,434

FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

CLASSIFICATION:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 91:

SEQUENCE CHARACTERISTICS:

LENGTH: 279 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: LUNGFE703

CLONE: 1234795

US-09-016-434-91

Query Match

Best Local Similarity 11.2%; Score 223; DB 4; Length 279;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGCCCATGAGATGTTGAACTCAATTTTACTGAGAGGTTGGGCTGCTGAGAA 60
DB 25 ATGGCCCATGAGATGTTGAACTCAATTTTACTGAGAGGTTGGGCTGCTGAGAA 84
```

QY 61 AGTGAACGGAAAGAGCTGCTAATTGATAGCGCGCATTTGGAGATACATACATCC 120
DB 85 AGTGAACGGAAAGAGCTGCTAATTGATAGCGCGCATTTGGAGATACATACATCC 144
QY 121 CACATTTGGAGAGCATTAATATCACTGCTCCAGCTTATGAGAGCGAGTTGCAACAG 180
DB 145 CACATTTGGAGAGCATTAATATCACTGCTCCAGCTTATGAGAGCGAGTTGCAACAG 204
QY 181 GACAAAGTGTAAATTACAGAGCTTCACAGCATTTACAGCGAAAC 223
DB 205 GACAAAGTGTAAATTACAGAGCTTCACAGCATTTACAGCGAAAC 247

RESULT 5
US-09-557-921-1
; Sequence 1, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125, 416
; CURRENT APPLICATION NUMBER: US/09/557, 921
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-557-921-1

Query Match 6.2%; Score 124.6; DB 4; Length 1830;
Best Local Similarity 56.7%; Pred. No. 7, 2e-31;
Matches 251; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

QY 461 TTGCCAATTTGGGCAACCCGAAATCTTCCCAATCTTATCTTGGCTGCAAGAGATG 520
DB 1036 TCGAGAACGCTGAGCTCACCCCACTTGGCTTCTGCTTGGCAATGACAGATG 1095
QY 521 TCCCAACAAGAGCTGATACAGACAGATGGAATGGTATGTTAAATGCACTATA 580
DB 1096 CTGAGACCTGAGACCATGACGCGCTGAACATCGCTACGTCATCACTACCACTC 1155
QY 581 CCGTCCAAAGCTGACTTT--ATCCCGAGTCTCATTTCTGCGTGGCTGTGAATG 637
DB 1156 ATCTCCCTCTACACTATGAGAAAGGCTGTTCAACTACAGCGGCTGCCAGCACTG 1215
QY 638 ACAGCTTTTGTGAGAAATTTGCGGTGTTGAGCAAAATCAGTAGATTTTATTGAGAAAG 697
DB 1216 ACAGCAACAAGCAAGCACTCGGCAAGTACTTTGAAAGGCTTTTGAAGTTCATTGAGGAAG 1275
QY 698 CAAGAAGCTCAATGATGATGTTCTAGTGCACTGTTTGAAGTGGAGATCTCCGCTCCGCCA 757
DB 1276 CTGACCAAGTGGAGAGGAGGCTTCTCATCTCACTGCGAGGCTGGGGTCTCCGCTCCGCCA 1335
QY 758 CCAATGCTATGCTTACATCATGAGAGATGACATGCTTTTATGATGAGAGTTACAGAT 817
DB 1336 CCAATGCTATGCTTACATCATGAGAGATGACATGCTTTTATGATGAGAGTTTAAAT 1395
QY 818 TTGTGAGAGAGAGAGAGCTTATATCTCAAACTTCAATTTTCTGGGCAACTCTG 877
DB 1396 TTGTGAGAGAGAGAGAGCTTATATCTCAAACTTCAATTTTCTGGGCAACTCTG 1455
QY 878 ACTATGAGAGAGATTAAGAAC 900
DB 1456 AGTTGAGAGAGACTTAACAC 1478

RESULT 6
US-09-922-146-3
; Sequence 3, Application US/09922146

; Patent No. 6566133
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowert
; APPLICANT: Brett P. Morla
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 9 EXPRESSION
; FILE REFERENCE: RTS-0252
; CURRENT APPLICATION NUMBER: US/09/922, 146
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 3
; LENGTH: 2303
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (114)...(1268)
US-09-922-146-3

Query Match 5.4%; Score 106.8; DB 4; Length 2303;
Best Local Similarity 55.3%; Pred. No. 8, 7e-25;
Matches 229; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

QY 475 CCAACCGAATCTTCCCAATCTTATCTTGGCTGCAAGAGATGCTCAACAGAG. 534
DB 723 CCGTCCAGATCTGCGCAACCTATCTGGGAGAGGCGGAGTTCCGCAATTTGGAG 782
QY 535 CTGATACAGCAAGATGGATGGTTATGTTTAAATGCCAGCTATACCTGTCGAA--G 591
DB 783 AGCCTGCCCAACATGGGCAATCCGCTACATCTCAATGTCACCCCAACCTCCAACTTC 842
QY 592 CCGACTTTATCCCGAGTCTCATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAG 851
DB 843 TTGAGAGAGATGATGATCTTCACTACAGAGATCCCATCTCCAGACCTGAGACGAG 902
QY 652 AAAATTTGCGGTGTTGAGCAAAATCAGATGATTTTATGAGAGAGCAAAAGCTTCAAT 711
DB 903 AACCTGCGGGTCTTTCGAGGCAATGATGATGATGATGAGGCTGTGTCAGAAC 962
QY 712 GATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771
DB 963 TCGGGGCTCTCTCACTGCTTGGGGGCTGAGCGGCTGTCACCGCTGCTGAGCC 1022
QY 772 TACATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 831
DB 1023 TACCTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1082
QY 832 AGACTTATATCTCAAACTTCAATTTTGGGCAACTCTGACTATGAG 885
DB 1083 AGCTTATATCTCCCAACTTCAATTTGAGGAGAGTGTGCTGAGCTTTGAG 1136

RESULT 7
US-09-016-434-1135
; Sequence 1135, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2


```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1135:
SEQUENCE CHARACTERISTICS:
LENGTH: 2109 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1418933
US-09-016-434-1135

Query Match
Best Local Similarity 4.8%; Score 96.2; DB 4; Length 2109;
Matches 222; Conservative 0; Mismatches 188; Indels 3; Gaps 1;

QY 484 ATCTTCCCAATCTTTATCTTGGCTGCGCAGAGATGTCCTCAACAAGAGCTGATACG 543
DB 979 ATCTTCCCTTCTCTACTTGGCTGCGCAAGACTCCACCACTTGACGCTTGAG 1038
QY 544 CAGATGGGATGGTGTATGTGTTAATGCCAGCTATACCTGTCCAAAGC---CTGACTT 600
DB 1039 GAATTCGCATCAAGTACATCTTGAAGCATCCCAATTTGCCAATCTCTTGAGAAC 1098
QY 601 ATCCCGAGTCTATTTCTGCGCTGCTGCTGAATGACAGCTTTTGAGAAAATTTTG 660
DB 1099 GCAGAGAGTTTAATCAAGCAAAATCCCATCTCGATCTGAGCCAAACCTGTCC 1158
QY 661 CCGTGTGGCAAAATCAGATGATTTCAATGAGAAAGAAAGCCCTCCATGATGATG 720
DB 1159 CAGTTTTCCCTGAGGCATTTCTTCAATGATGAGAGCCGGGCAAGAACTGTGATC 1218
QY 721 CTAGTGACCTGTTTAACTGCGATCTCCGCTCCGCAACATGCTATGCTATCATG 780
DB 1219 TTGTATCATTTGCTGGCTGCAATGACCGCTCAGTCACTGTGCTTACCTTATG 1278
QY 781 AAGAGATGAGATGCTTTAGATGAGCTTACAGATTTGTGAAAAGAAAAGAACTACT 840
DB 1279 CAGAACTCAATCTGTGATGAGATGCTATGACATTTGCAAAATGAAAATCCAC 1338
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCCGAGCATATGAGAAAGAT 893
DB 1339 ATATCCCTCAACTCACTTCAATGGGTCAAGTCTGAGACTTCAGAGAGAGCT 1391

RESULT 8
US-09-023-655-946
; Sequence 946 Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
```

```

CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: US/09/023,655
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 946:
SEQUENCE CHARACTERISTICS:
LENGTH: 2109 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1418933
US-09-023-655-946

Query Match
Best Local Similarity 4.8%; Score 96.2; DB 4; Length 2109;
Matches 222; Conservative 0; Mismatches 188; Indels 3; Gaps 1;

QY 484 ATCTTCCCAATCTTTATCTTGGCTGCGCAGAGATGTCCTCAACAAGAGCTGATACG 543
DB 979 ATCTTCCCTTCTCTACTTGGCTGCGCAAGACTCCACCACTTGACGCTTGAG 1038
QY 544 CAGATGGGATGGTGTATGTGTTAATGCCAGCTATACCTGTCCAAAGC---CTGACTT 600
DB 1039 GAATTCGCATCAAGTACATCTTGAAGCATCCCAATTTGCCAATCTCTTGAGAAC 1098
QY 601 ATCCCGAGTCTATTTCTGCGCTGCTGCTGAATGACAGCTTTTGAGAAAATTTTG 660
DB 1099 GCAGAGAGTTTAATCAAGCAAAATCCCATCTCGATCTGAGCCAAACCTGTCC 1158
QY 661 CCGTGTGGCAAAATCAGATGATTTCAATGAGAAAGAAAGCCCTCCATGATGATG 720
DB 1159 CAGTTTTCCCTGAGGCATTTCTTCAATGATGAGAGCCGGGCAAGAACTGTGATC 1218
QY 721 CTAGTGACCTGTTTAACTGCGATCTCCGCTCCGCAACATGCTATGCTATCATG 780
DB 1219 TTGTATCATTTGCTGGCTGCAATGACCGCTCAGTCACTGTGCTTACCTTATG 1278
QY 781 AAGAGATGAGATGCTTTAGATGAGCTTACAGATTTGTGAAAAGAAAAGAACTACT 840
DB 1279 CAGAACTCAATCTGTGATGAGATGCTATGACATTTGCAAAATGAAAATCCAC 1338
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCCGAGCATATGAGAAAGAT 893
DB 1339 ATATCCCTCAACTCACTTCAATGGGTCAAGTCTGAGACTTCAGAGAGAGCT 1391

RESULT 9
US-09-023-655-347
; Sequence 347 Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
```

APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 347:
SEQUENCE CHARACTERISTICS:
LENGTH: 1208 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THYRNOT03
CLONE: 1444245
US-09-023-655-347

Query Match 4.7%; Score 94.6; DB 4; Length 1208;
Best Local Similarity 64.3%; Pred. No. 6.7e-21;
Matches 142; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 680 TAGATTTCATTGGAAGCAAAAGCCCAATGGATGTTCTAGTGCATCTGTTAGCTG 739
DB 4 TTGATTCATTGGAAGCAAGCTCAAGTGTGGAAAGGGGCTTCATCCTGCGAGGCTG 63
QY 740 GGATCTCCGCTCCGCCACCATGCTATCGCTACATCATGAGAGAGATGATCTT 799
DB 64 GGGTCTCCGCTCCGCCACCATGCTATCGCTTACTGATGGAAGCAACATCGATGCCA 123
QY 800 TAGATGAAGCTTACAGATTGTTGTAAGAAAAAGAAAGAAAGAAAGAAAGAAAG 859
DB 124 TGATGATGCTTATTAATTTGTCAAAAGCAAGCAACCAATTAATCTCCCAACTTA 183
QY 860 TTCGGGGCAACCTCGGACATGATGAGAAAGAAAGATTAGAAC 900
DB 184 TCATGGGGCAGTGTGATGAGTTCGAGGAGAACTTAACAC 224

RESULT 10
US-09-702-705-801
Sequence 801, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 801
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapien
US-09-702-705-801

Query Match 4.7%; Score 93.6; DB 4; Length 1619;
Best Local Similarity 52.3%; Pred. No. 1.8e-20;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

QY 484 ATTCTTCCCAATCTTATCTTGCTGCCAGCGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTTCTTACCTCGGCACTGCTACATGCTGCCGAGAGACATGCTGAC 379
QY 544 CAGATGGATGCTTATGTTAATGCCAGTAACTGTCGAAAGCTGACTTATC 603
DB 380 GCCCTGGGATCAGGCTCTGTTGAATGTCCTCGGACTGCCCAAC--ACTTGA 436
QY 604 CCGAGTCTATTCCTGGGTGCTGCTGGAATGACAGCTTTGTGAAATTTGCCG 663
DB 437 GCACTATGATGACAGTCAATCCAGTGAAGATACCAAGGCCGACATCAGCTC 436
QY 664 TGGTTGACAAATCAGTAAATTTTATTGAGAAAGAAAGCCCTCAATGATGTTCTA 723
DB 497 TGTTTATGGAAGCCATAGATCATCATGATCGGTGAAGAGCTGCGGCGGCTG 556
QY 724 GTGACTGTTTAACTGGGATCTCCGCTCGCCACATGCTATCGCTTACATGAG 783
DB 557 GTGACTGCGAGCGGAGCATCTCGCGTGGCCACCATCTGCTGCTTACCTGATGATG 616
QY 784 AGATGACATGCTTTAGATGAGCTTACAGATTGTGAAAGAAAAAGACTTACTAT 843
DB 617 AAGAAACGGGTAGGCTGAGAGGCTTGAAGTGTGAAGAGCGCGGACATTAATC 676
QY 844 TCTCAAACTTCAATTTCTGGGCAACTCTGAGCTATGAGAAAGAAATTAAAGACAG 903
DB 677 TGCCCACTTCACTTATGAGGCGAGCTGCTGCAATGAGTCCAGGTGCTGCGACG 736
QY 904 ACTGAGCATCAGGGCCAAAGAC 927
DB 737 TCCTGTGCTGCGAGGCTGCTAGC 760

RESULT 11
US-09-736-457-801
Sequence 801, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 801
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-801

Query Match
Best Local Similarity 4.7%; Score 93.6; DB 4; Length 1619;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

```
QY 484 ATCTTCCCATCTTTATCTTGGCTGCGACGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCATCTTCTCTACCTCGGACGCTTACCATGCTGCGGAGAGACATGCTGAC 379
QY 544 CAGAAATGGGATTTGGTTATGTTAAATGCCCAAGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGGACATCAGGCTCTGTTGAATGTCCTCGGACTGCCAAAC---ACTTTGAA 436
QY 604 CCCGAGTTCATTTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
DB 437 GGAAGTATCAGTACAAAGTGCATCCCAAGTGAAGTAAACAAGGCGGACATCAGCTCC 496
QY 664 TGGTTGGACAAATTCAGTATTTCAATTGAGAAAGCAAAAGCTCCAAATGATGTTCTA 723
DB 497 TGGTTATGAGAACATAGATAGATGATGATGATGATGATGATGATGATGATGATGATG 723
QY 724 GTGCACTGTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 783
DB 557 GTGCACTGCTGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
QY 784 AGATGAGCAATGCTTTTATGATGAGCTTACAGATTTGTGAAGAAAGAAAGAAAGCT 843
DB 617 AAGAAAGGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 676
QY 844 TCTCCAACTTCAATTTCTGAGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 903
DB 677 TCGCCCACTTCAAGCTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 736
QY 904 ACTGAGCATCAGGCGCAAGAGC 927
DB 737 TCTGTGCTGCGAGGCTGCTGAC 760
```

RESULT 12

US-09-614-124B-801
Sequence 801, Application US/09614124B
Patent No. 6630574
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.478C9
CURRENT APPLICATION NUMBER: US/09/614,124B
CURRENT FILING DATE: 2001-07-11
NUMBER OF SEQ ID NOS: 1668
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 801
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapien
US-09-614-124B-801

Query Match
Best Local Similarity 4.7%; Score 93.6; DB 4; Length 1619;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

```
QY 484 ATCTTCCCATCTTTATCTTGGCTGCGACGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCATCTTCTCTACCTCGGACGCTTACCATGCTGCGGAGAGACATGCTGAC 379
QY 544 CAGAAATGGGATTTGGTTATGTTAAATGCCCAAGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGGACATCAGGCTCTGTTGAATGTCCTCGGACTGCCAAAC---ACTTTGAA 436
QY 604 CCCGAGTTCATTTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
DB 437 GGAAGTATCAGTACAAAGTGCATCCCAAGTGAAGTAAACAAGGCGGACATCAGCTCC 496
QY 664 TGGTTGGACAAATTCAGTATTTCAATTGAGAAAGCAAAAGCTCCAAATGATGTTCTA 723
DB 497 TGGTTATGAGAACATAGATAGATGATGATGATGATGATGATGATGATGATGATGATG 723
QY 724 GTGCACTGTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 783
DB 557 GTGCACTGCTGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
QY 784 AGATGAGCAATGCTTTTATGATGAGCTTACAGATTTGTGAAGAAAGAAAGAAAGCT 843
DB 617 AAGAAAGGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 676
QY 844 TCTCCAACTTCAATTTCTGAGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 903
DB 677 TCGCCCACTTCAAGCTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 736
QY 904 ACTGAGCATCAGGCGCAAGAGC 927
DB 737 TCTGTGCTGCGAGGCTGCTGAC 760
```

RESULT 13

US-09-671-325-801
Sequence 801, Application US/09671325
Patent No. 6667154
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C12
CURRENT APPLICATION NUMBER: US/09/671,325
CURRENT FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 1825
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 801
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapien
US-09-671-325-801

Query Match
Best Local Similarity 4.7%; Score 93.6; DB 4; Length 1619;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

```
QY 484 ATCTTCCCATCTTTATCTTGGCTGCGACGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCATCTTCTCTACCTCGGACGCTTACCATGCTGCGGAGAGACATGCTGAC 379
```

QY 544 CAGATGGAGTGTATATGTTAAATGCCAGCTATACCTGCCAAAGCCTGACTTATC 603
DB 380 GCCCTGGGATCAGGGCTCTGTTGAATGTCTCTCGAGACTGCCAAAC--ACTTTGAA 436
QY 604 CCGGAGTCTATTTCTCGCTGTGCTGTGAATGACAGCTTTTGTGAAAATTTTCCG 663
DB 437 GGACATATACAGTAAAGTGCATCCGAGTGGAAGATACCAAGGCCAGCATCAGCTCC 496
QY 664 TGGTTGACAAATCAGTAAATTTTCTTGAAGAAAAGCAAAACCTTCATGATGTCTTA 723
DB 497 TGGTTACAGAACCCATAGATCATGATGCGGTGAAGAGATGCGCTGGGCGGTGCTG 556
QY 724 GTCACTGTTTAAAGTGGGATCTCCGCTCCGCCACCATCTGCTATGCGCTCATATGAA 783
DB 557 GTGCACTGCCAGGCGGGCATCTGCGGTCCGCCACCATCTGCTGCTTACCTGATGATG 616
QY 784 AGATGACATGTCTTTAATGATGAGCTTACAGATTTGTGAAGAAAAGAACTACTATA 843
DB 617 AAGAAAGGGGTGAGGCTGAGAGAGGCTTCGAGTTGTTAAGAGCGCGCAGCATTATC 676
QY 844 TCTCCAACTTCAATTTTCTGGGCCCACTCTGACATATAGAAAGATTAAGAACAG 903
DB 677 TCGCCCAACTTCACTTCAATGGGGCAGCTGCTCAGATTCCAGGTGCTGCGCACG 736
QY 904 ACTGAGCATCAGGGCCAAAGAGC 927
DB 737 TCCTGTCTGCGGAGGCTGCTAGC 760

RESULT 14

US-09-589-184-801
; Sequence 801, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaltanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589.184
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-589-184-801

Query Match 4.7%; Score 93.6; DB 4; Length 1619;
Best Local Similarity 52.3%; Pred. No. 1.8e-20;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

QY 484 ATCTTCCCAATCTTATCTTGGCTGCCAGAGATGTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTCTCTTACTCTGCGAGTGTCTACCATGTCTCCGGAGACATGTGAGAC 379
QY 544 CAGATGGAGTGTATATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGGATCAGGGCTCTGTTGAATGTCTCTCGGAGCTGCCAAAC--ACTTTGAA 436
QY 604 CCGGAGTCTATTTCTCGCTGTGCTGTGAATGACAGCTTTTGTGAAAATTTTCCG 663
DB 437 GGACATATACAGTAAAGTGCATCCGAGTGGAAGATACCAAGGCCAGCATCAGCTCC 496
QY 664 TGGTTGACAAATCAGTAAATTTTCTTGAAGAAAAGCAAAACCTTCATGATGTCTTA 723
DB 497 TGGTTACAGAACCCATAGATCATGATGCGGTGAAGAGATGCGCTGGGCGGTGCTG 556
QY 724 GTCACTGTTTAAAGTGGGATCTCCGCTCCGCCACCATCTGCTATGCGCTCATATGAA 783
DB 557 GTGCACTGCCAGGCGGGCATCTGCGGTCCGCCACCATCTGCTGCTTACCTGATGATG 616
QY 784 AGATGACATGTCTTTAATGATGAGCTTACAGATTTGTGAAGAAAAGAACTACTATA 843
DB 617 AAGAAAGGGGTGAGGCTGAGAGAGGCTTCGAGTTGTTAAGAGCGCGCAGCATTATC 676

DB 497 TGGTTACAGAACCCATAGATCATGATGCGGTGAAGAGATGCGGTGGGCGGTGCTG 556
QY 724 GTGCACTGTTTAAAGTGGGATCTCCGCTCCGCCACCATGCTTATGCGCTCATATGAA 783
DB 557 GTGCACTGCCAGGCGGGCATCTGCGGTCCGCCACCATGCTGCTGCTTACCTGATGATG 616
QY 784 AGATGACATGTCTTTAATGATGAGCTTACAGATTTGTGAAGAAAAGAACTACTATA 843
DB 617 AAGAAAGGGGTGAGGCTGAGAGAGGCTTCGAGTTGTTAAGAGCGCGCAGCATTATC 676
QY 844 TCTCCAACTTCAATTTTCTGGGCCCACTCTGACATATAGAAAGATTAAGAACAG 903
DB 677 TCGCCCAACTTCACTTCAATGGGGCAGCTGCTCAGATTCCAGGTGCTGCGCACG 736
QY 904 ACTGAGCATCAGGGCCAAAGAGC 927
DB 737 TCCTGTCTGCGGAGGCTGCTAGC 760

RESULT 15

US-09-702-705-804
; Sequence 804, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaltanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702.705
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 804
; LENGTH: 4637
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-804

Query Match 4.7%; Score 93.6; DB 4; Length 4637;
Best Local Similarity 52.3%; Pred. No. 4.3e-20;
Matches 232; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

QY 484 ATCTTCCCAATCTTATCTTGGCTGCCAGAGATGTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTCTCTTACTCTGCGAGTGTCTACCATGTCTCCGGAGACATGTGAGAC 379
QY 544 CAGATGGAGTGTATATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGGATCAGGGCTCTGTTGAATGTCTCTCGAGCTGCCAAAC--ACTTTGAA 436
QY 604 CCGGAGTCTATTTCTCGCTGTGCTGTGAATGACAGCTTTTGTGAAAATTTTCCG 663
DB 437 GGACATATACAGTAAAGTGCATCCGAGTGGAAGATACCAAGGCCAGCATCAGCTCC 496
QY 664 TGGTTGACAAATCAGTAAATTTTCTTGAAGAAAAGCAAAACCTTCATGATGTCTTA 723
DB 497 TGGTTACAGAACCCATAGATCATGATGCGGTGAAGAGATGCGCTGGGCGGTGCTG 556
QY 724 GTCACTGTTTAAAGTGGGATCTCCGCTCCGCCACCATGCTATGCGCTCATATGAA 783
DB 557 GTGCACTGCCAGGCGGGCATCTGCGGTCCGCCACCATGCTGCTTACCTGATGATG 616
QY 784 AGATGACATGTCTTTAATGATGAGCTTACAGATTTGTGAAGAAAAGAACTACTATA 843
DB 617 AAGAAAGGGGTGAGGCTGAGAGAGGCTTCGAGTTGTTAAGAGCGCGCAGCATTATC 676

Oy 844 TCTCCAAACTTCATTTTCTGGGCAACTCTGACTATGAGAGAATTAAAGACAG 903
Db 677 TGGCCAACTTCAGCTTCATGGGGCAGCTGCTGAGTCCAGGTCTGGCCAAG 736
Oy 904 ACTGAGCATCAGGGGCCAAGAGC 927
Db 737 TCTGTGCTGCGGAGGCTGCTAGC 760

Search completed: June 22, 2004, 02:56:08
Job time : 168.047 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_n2p model

Run on: June 21, 2004, 12:28:21 ; Search time 95.5895 Seconds
(without alignments)
11793.817 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 3615
Sequence: 1 atggcccatgagatgattg.....tggaaatcattgaggtctcc 1995

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=xjp
-Q/cgnt 1/USPTO.spool.p/US10029345/runat.21062004.122814.4167/app.query.fasta_1.3278
-DB=A.Geneseq.29345a-108.fasta -SUFFIX=rag -MINMATCH=0.1 -LOOPTOL=0
-LOPEXT=0 -INITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US10029345 @CGN 1.128 @runat.21062004.122814.4167 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: A_Geneseq.29345a-108.fasta
1: geneseqp1980a:*
2: geneseqp1990a:*
3: geneseqp2000a:*
4: geneseqp2001a:*
5: geneseqp2002a:*
6: geneseqp2003a:*
7: geneseqp2003bs:*
8: geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	3418	94.6	665 5	ABR52381 Protein x
2	3418	94.6	665 5	ABR52407 Protein x
3	3406	94.2	665 4	AAE04834 Human SGP
4	3406	94.2	665 4	AAU09016 Human dua
5	3406	94.2	665 5	AAU09156 Human dua
6	3406	94.2	665 5	AAU09946 Protein s
7	3406	94.2	665 5	AAU75789 Human pro
8	3406	94.2	665 5	ABR97946 Human pro
9	3406	94.2	665 5	AAU79929 Human dua
10	3406	94.2	665 5	ABR97291 Novel hum

11	3399	94.0	665 6	ADA54744 Human pro
12	3398	94.0	665 5	AAU79161 Human dua
13	3396	93.9	665 5	AAU79162 Human dua
14	3388	93.7	665 5	ABR52352 Protein x
15	3379.5	93.5	666 4	ABR520325
16	3368.5	93.2	664 5	ABR52424
17	3322	91.9	672 4	AAU52744 Human pro
18	3057.5	84.6	660 5	ABR52385 Protein x
19	2594	71.8	517 5	AAU79159 Human dua
20	1542	42.7	302 5	ABR52425
21	1326	36.7	625 5	ABR52382 Protein x
22	1326	36.7	625 5	ABR52350 Protein x
23	1326	36.7	625 6	ABG73440 Human dua
24	1302	36.0	663 5	ABR52351 Protein x
25	1297	35.9	663 2	AAU92150
26	977.5	27.0	579 7	ADG08458
27	742.5	20.5	375 4	ABG07902 Novel hum
28	741	20.5	140 5	ABR52404
29	716	19.8	140 5	ABR52405
30	654.5	18.1	170 4	AAE66436
31	654.5	18.1	170 4	AAE06780
32	476	13.2	482 4	AAU73225
33	471.5	13.0	381 5	ABR52383
34	471.5	13.0	381 6	ABP6803
35	469	13.0	444 5	AAO20515
36	469	13.0	482 3	AAU18655
37	469	13.0	482 3	AAE29641
38	469	13.0	482 5	ABR52384
39	469	13.0	482 5	ABJ05600
40	466.5	12.9	381 6	ABP57087
41	465.5	12.9	381 7	ADD48300 Rat Prote
42	465.5	12.9	381 7	ADG62625
43	455	12.6	394 4	AAU76875
44	455	12.6	394 5	AAU85530
45	455	12.6	394 6	ABU69502 Humna STY

ALIGNMENTS

RESULT 1	ID	ABR52381	standard; protein; 665 AA.
AC	XX	ABR52381,	
DT	XX	19-JUN-2003	(first entry)
DE	XX	Protein relating to the invention SEQ ID NO: 109.	
KW	XX	antiproliferative; hepatotropic; nephrotropic; antiarthritic;	
KW	XX	antiproliferative; cardiatic; cytostatic; gene therapy; liver disease;	
KW	XX	proliferative disorder; renal failure; cardiovascular disorder;	
KW	XX	immunological disorder; arthritis; psoriasis; congenital heart defect;	
KW	XX	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
OS	XX	Homo sapiens.	
PN	XX	WO200257460-A2.	
PD	XX	25-JUL-2002.	
PF	XX	20-DEC-2001, 2001MO-US050459.	
PR	XX	20-DEC-2001, 2000US-0256668P.	
PR	XX	30-MAR-2001, 2001US-0280186P.	
PR	XX	01-MAY-2001, 2001US-0287735P.	
PR	XX	05-JUN-2001, 2001US-0295848P.	
PR	XX	25-JUN-2001, 2001US-0300465P.	
PA	XX	(BRIM) BRISTOL-MYERS SQUIBB CO.	
PI	XX	Jackson DG, Feder J, Nelson T, Muntier G, Ramanathan C, Lee L,	
PI	XX	Stiemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,	

QY	1621	TOGGATATCTGGGCCCCCAGACGCTTACGCCCTTCCCGACAGACAGCTGATTTGGC	1680
Db	541	SeraspIleuAlaProGlnThrSerThrProSerLeuIlnrSerSerIrrPheAla	560
QY	1681	ACAGAGTCTCTACACTTCTACTCTGTCCCTCAGCCCACTTACGGAGGACAGTGCAGTTACTCT	1740
Db	561	ThrGlnSerSerHisPheTyrSerHisSerAlaSerIleIleTyrGlyGlySerAlaSerTyrSer	580
QY	1741	GCTTACAGCTGCAGCAGCTGCCCCCACTTGGCGAGACCAAGTCTAATTCTGTGCGCAGGCGG	1800
Db	581	AlaTyrSerCysSerGlnLeuProThrCysGlyIlePheGlnValTyrSerValArgArg	600
QY	1801	CAGAAAGCCAAATGATACAGAGCTGACCTCGGGGGCGAGCTGGCATGAAAGAGAGGCCCTTTGAA	1860
Db	601	GlnIlypProSerAspArgAlaAspSerTrgArGserTrpHisGlnGlnIlySerProPheGln	620
QY	1861	AAGCAGTTTAACGCGAGAGAGCTGCGCAATGTGAATTGTGAGAGACATCATCTCAGAGAAC	1920
Db	621	LysGlnPheIlyAsnArgSerCysGlnMetGlnPheGlyIlySerIleMetSerGlnAsn	640
QY	1921	AGGTTCACGGGAGAGAGCTGGGGAGAAAGTGGGACAGTCTAGCTTTTGGGCGAGCATGAA	1980
Db	641	ArgSerArgGlnGlnIleuGlyIlyValGlySerGlnSerSerPheSerGlySerMetGln	660
QY	1981	ATCATTTAGAGTCTCC	1995
Db	661	IleIleGlnValSer	665
RESULT 2			
ABRS52407			
ID	ABRS52407	standard; protein; 665 AA.	
XX	ABRS2407;		
XX	19-JUN-2003	(first entry)	
XX	DE	Protein relating to the invention SEQ ID NO: 148.	
XX	XX		
XX	XX	anticiproliferative; hepatotropic; nephrotropic; antiarthritic;	
XX	XX	antipneumatic; cardiac; cytotoxic; gene therapy; liver disease;	
XX	XX	proliferative disorder; renal failure; cardiovascular disorder;	
XX	XX	immunological disorder; arthritis; psoriasis; congenital heart defect;	
XX	XX	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
OS	Homo sapiens.		
XX	XX		
XX	XX	WO200257460-A2.	
XX	XX	25-JUN-2002.	
XX	XX	20-DEC-2001; 2001WO-US050459.	
XX	XX	20-DEC-2000; 2000US-0256868P.	
XX	XX	30-MAR-2001; 2001US-0280186P.	
XX	XX	01-MAY-2001; 2001US-0287735P.	
XX	XX	05-JUN-2001; 2001US-0295848P.	
XX	XX	25-JUN-2001; 2001US-0300465P.	
XX	XX	(BRIM) BRISTOL-MYERS SQUIBB CO.	
XX	XX	Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,	
XX	XX	Slimer N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,	
XX	XX	Krystek S, Mcatee P, Suchard S, Banas D;	
XX	XX	WPI; 599721/64.	
XX	XX	N-PSDB; ACC60572.	
XX	XX	Novel polynucleotides encoding human phosphatase polypeptides useful in	
XX	XX	the prevention or treatment of e.g. proliferative and cardiovascular	
XX	XX	disorders.	
XX	XX	Disclosure; Fig 19; 801p; English.	

CC	The invention relates to a novel isolated nucleic acid comprising a
CC	polynucleotide having a nucleotide sequence selected from 40
CC	polynucleotides fully defined in the specification. The polynucleotide of
CC	the invention has antiproliferative, hepatotropic, nephrotropic,
CC	antiarthritic, antisporadic, cardiatic, and cytostatic activity. The
CC	polynucleotide may have a use in gene therapy. A polynucleotide or
CC	polypeptide of the invention is useful for preventing, treating or
CC	ameliorating a medical condition, e.g. a proliferative disorder. They are
CC	also useful for treating e.g. liver disease, renal failure, immunological
CC	disorders including arthritis and psoriasis, cardiovascular disorders
CC	such as congenital heart defects and congestive heart failure, and
CC	cancer. A method of the invention is useful for diagnosing a pathological
CC	condition or susceptibility to a condition in a subject. The present
CC	sequence is used in the exemplification of the invention
SQ	Sequence 665 AA:
Alignment Scores:	
Pred. No.: 2,85e-288	Length: 665
Score: 3418.00	Matches: 665
Percent Similarity: 100.00%	Conservative: 0
Best Local Similarity: 100.00%	Mismatches: 0
Query Match: 94.55%	Indels: 0
DB: 5	Gaps: 0
US-10-029-345A-108_COPY 538_2532 (1-1995) x ABR52407 (1-665)	
OY	1 ATGGGCCATGAGATCATTTGGAACCTCAAAATTGTACTGAGAAGTGGTGCGCTGCTGAA 60
Dd	1 MetAlahIsGlumEtIlEgIyThrInIlleValIThrGIuArgLeuValAlalaLeuEnGLu 20
OY	61 AGTGAACGGAAGAAAAGTCGTGTAATTGANTAGCCGCGCATTGTGTGGAATAACAATACATCC 120
Dd	21 SerGIyThrGIuLyVallLeuEnlIeaPseRarProPhValGIurYanThrSer 40
OY	121 CACATTTTGGAAAGCATTAATATCACTGCCTCCAAAGCTTATGGAAGGAAGTTGCAACAG 180
Dd	41 HisIIeuNGluAlaIlaAsnIIeaSnlCySeRIySleMeLlyArGIArgLeuGIingIn 60
OY	181 GACAAAGGTAAATATCACAGAGCTCATCCAGACATTCAGCAAACATPAAGGTGACATTGAT 240
Dd	61 AspyIValIleuIIethrGIuEnlIegInHIsSerIaLyHisIlyValaPleIeaP 80
OY	241 TGCAGTCGAAGGTTGATGATTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTCA 300
Dd	81 CyserGIuLyValValValITyrAbpGINserSerGIaNPvalAlaSerIeusSer 100
OY	301 GACTGTTTCTCACTGTACTTCTGGGTAACTGGAAGAAGCTTCACTCTGTTCACCTG 360
Dd	101 AspCySPheLeuThrValIleuEnlIGlyLSerPheXnsErValIHISleu 120
OY	361 CTTCAGAGTGGGTTTCTGAGATTCCTCGTGGTTCCCTGGGCTCTGGAAGAAATCC 420
Dd	121 LeuAlaGIyLIypHeaIaGIuPheSerxArgCPheProGIyLeuCYbGIuGIyLSer 140
OY	421 ACTAGATCCCTACCTGCATTTCTCAGCCTTGCTTACCTGTGGCCAATTGGGCAACC 480
Dd	141 ThrIeuValIProTHrCySIleSerGIuProCySleuProValAlaSnIIeGIyProTHr 160
OY	481 CGAATTCCTCCCAATCTTTATCTTGCGTCCAGCGAGATGTCCTCAACAAGAGCTGATA 540
Dd	161 ArgIIeuNPProAnLeuTyfLeuGIyCySGInIaGaPvalIleuAnsnlybGIuEnlIe 180
OY	541 CAGCAAGATGGAGTGGTATATGTGTTAAATGCCACTATACCTGTGCCAAAGCTGACTT 600
Dd	181 GIuGIInsnGIyIIeGIyTyrValIleuAnsnIaSerTYrTHrCySProLySPraSPpe 200
OY	601 ATCCCGAGTCTCATTTCTCGCGTGTGCTGAGATGACAGCTTTGTGAGAAAAATTGG 660
Dd	201 IlIeproGIuSerThIspHeuArgValIProValaAnsnPseRPheCySbGIuLySIleu 220
OY	661 CCGAGGTGGAACAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCCAATGATGTGTT 720

D	221	ProTPrleuAspIysSerValAspPheIleGluValAlaIysAlaSerAsnGlyCysVal	240
Q	721	CTAGTGCACCTGTTTAGCTGGGATCTCCCGCTCCGACCAATCGCTAATGCGCTTACATCAG	780
D	241	LeuValHisCysLeuAlaGlyIleSerTrpSerAlaThrIleAlaIleAlaTyIleMet	260
Q	781	AAGAGAAATGAGCATCTCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAGAAACACTACT	840
D	261	LysAlaGluAspMetSerLeuAspGluAlaTyTrpPheValIysGluLysArgProThr	280
Q	841	ATATCTCCAAATTCATTTCTGGGGCAATCCCTGGACATATAGAAAGAAATTAAAGAAC	900
D	281	IleSerProAsnPheAsnPheIleGlyGlnLeuLeuAspTyrGluLysIleLysAsn	300
Q	901	CAGACTGGAGCATAGAGGCCAAAGAGCAAACTCAAGCTGTGCACCTGGAGAAAGCCAAAT	960
D	301	GlnThrGluAlaSerGlyProLysSerIysLeuLysLeuHisIleGluLysTrpAsn	320
Q	961	GAACCTGTCCCTGTGTCTCAGAGGGGTGACAAAGAGAGAACGCCCTCATGTCACCC	1020
D	321	GlnProValProAlaValaSerGluGlyGlnLysSerGluThrProLeuSerProPro	340
Q	1021	TGAGCGACCTCTCTCTCAGAGGACAGAGCAAAAGCCCGTGCATCCCGCAGGCTG	1080
D	341	CysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSerVal	360
Q	1081	CCGACGGTCCCGACCGGTGAGCCGTGCGTTAGAGGACACCCCGCTGTACAGCGCTC	1140
D	361	ProSerValProSerIleGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu	380
Q	1141	AGTGGCTCTCACCTCTCCGACAGACAGGCTGGAAGACAGCAATAGCTCAAGCGTCCCTTC	1200
D	381	SerGlyLeuHisIleSerSerAlaAspArgLeuGlnAspSerAsnLysLeuLysIleSerPhe	400
Q	1201	TCTTGGATATCAAACTCAATTCAATTCACGSCAGCATGGCAGCATCTTACATGGCTTC	1260
D	401	SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe	420
Q	1261	TCCATATCAGAAAGATGCTTTGGAATCTTACAAACCTTCCATCTCTGGATGGAGCAAC	1320
D	421	SerSerSerGluAspAlaIleGluGlyTyrTyrLysProSerThrThrLeuAspGlyThrAsn	440
Q	1321	AAGCTATGCCAGTTCTCCCTGTTACAGAACTTGGAGACATGCCAGATCCGAAACAGTCTT	1380
D	441	LysLeuLysGlnPheSerProValGlnGlnLeuSerGlnThrProGlnThrSerPro	460
Q	1381	GATTAAGAGAGAGCGACATCCCCAGAAAGACTTCAGACCCGACGCTTTCAGACGCCAG	1440
D	461	AspLysGlnGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln	480
Q	1441	AGCAAGGATTCGATTCCGTCAGAACACAGACAGATGGACACGCCCGCAGAGTCCCTTTTA	1500
D	481	SerTyrArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnHisSerLeuLeu	500
Q	1501	TCTCCACTGCATCGAAGTGGAGCGGTGAGAGCAATTACACACAGACTTCCTTTGGC	1560
D	501	SerProLeuHisArgSerGlySerValGlnAspAsnTyrHisThrSerPheLeuPheGly	520
Q	1561	CTTTCACCAAGCAGACAGCACTTCAGAAATCTGCTGCCCTGGAGCTTAAAGGCTGAC	1620
D	521	LeuSerThrSerGlnGlnHisIleThrIlySerAlaGlyLeuGlyLeuLysGlyTyrHis	540
Q	1621	TCGATATCTTGGCCCCCAACCTCTTACCCCTTCCCTGACACAGACTGTATTTTGGC	1680
D	541	SerAspIleLeuAlaProGlnThrSerThrProSerIleuThrSerSerThrPyrPheAla	560
Q	1681	ACAGATCTTCACATTTCTACTCTGCTGACCAATCAGAGAGCAGTGCACATTACTCT	1740
D	561	ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAlaSerTyrSer	580
Q	1741	GCTTACAGCTGACGCCAGCTGCCCACTTGGAGAGACCAAGTCAATCTGTGTGGCAGGGG	1800
D	581	AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgTrpArg	600

QY	1801	CGAAGCCCAAGTGAAGAGCTGACTGCGCGCGAGCTGGCATGGAAGAGCCCTTTGAA	1860
Db	601	GINLVSPPROSERASPARGALAApSERARgARgserTRpHISGLUGLuserPRPpHeLU	620
QY	1861	AAGCAGTTTAAACGAGAAAGCTGCGCAAAATGGAAATTGGAGAGACATCATGTCAGAGAAC	1920
Db	621	LysGInpHeLysARgARgserCYGGLmECInpHeGLysuSerILlEmetSerGLuasn	640
QY	1921	AGGTACGCGGAAAGCTTCGCGGAAAAGTGGCAGTCACTTCTTTTCGGCGCAGATGGA	1980
Db	641	ARgserARgGLUGLunpHeuGLysVALGLysERGLnserSerPHeSerGLysERmetGLU	660
QY	1981	ATCATTGAGGCTCTCC	1995
Db	661	ILleILleGLuVALser	665
RESULT 3			
ID	AAE04834	standard; protein; 665 AA.	
AC	AAE04834;		
XX			
DT	10-SEP-2001	(first entry)	
XX			
DE		Human SGP002 phosphatase polypeptide.	
KW		Human; SGP002 phosphatase polypeptide; phosphatase-related disease;	
KW		immune-related disorder; ocular disease; organ transplant rejection;	
KW		infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;	
KW		metabolic disorder; haematopoietic cancer; mood disorder; cardiac;	
KW		Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;	
KW		cardiovascular disease; brain; neuronal-associated disease; dyskinesia;	
KW		attention disorder; cognitive disorder; psychiatric disorder; cytoskeletal;	
KW		neurological disorder; vincristine; nocrotropic; cerebroprotective; therapy;	
KW		neuroprotective; antibacterial; vulnertary; triamthiliser; antiaclastmatic;	
KW		hypertensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;	
KW		antifungal; dual specifically phosphatase; DSP; MAP kinase phosphatase;	
KW		MKP; migraine; chromosome 12p11.1-p12.1.	
XX			
OS		Homo sapiens.	
XX			
EH		Key	
FT		Location/Qualifiers	
FT		1. 173	
FT		/label= Catalytic domain	
FT		158. 297	
FT		/label= Phosphatase_domain	
XX			
PN		MO200146394-A2.	
XX			
PD		28-JUN-2001.	
XX			
PF		21-DEC-2000; 2000WO-US034736.	
XX			
PR		21-DEC-1999; 99US-0172255P.	
PR		28-DEC-1999; 99US-0175766P.	
PR		25-JUN-2000; 2000US-0178078P.	
PR		31-JUN-2000; 2000US-0179301P.	
PA		(SUGEN-) SUGEN INC.	
XX			
PI		Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;	
PI		Hill RJ, Flanagan P;	
XX			
DR		WPI: 2001-418058/44.	
XX		N-PESDB; AAD09492.	
PT		Novel phosphatase polypeptide useful for treating cancers, immune-related	
PT		diseases and disorders, cardiovascular disease, brain or neuronal-	
XX		associated diseases and metabolic disorders.	
PS		Claim 7; Fig 2; 1866p; English.	
XX			

QY 1741 GCCTACAGCTGAGCAGCTGCTCCCACTTCCGAGACCAAGCTATTCTGTGCGCAGCGG 1800
DB 581 ATATYRserCySeSerGlnneuProthrcysGlyaspGlnValTyrSerValArgArgArg 600
QY 1801 CAGAACCCAGTGAAGAGCTGACTCGCGCGAGCTGGAGCTGAAGAGCCCTTTGAA 1860
DB 601 GlnLysPProSerAspArgAlaAspSerArgArgSerTrpHisGlnGlnSerProPheGln 620
QY 1861 AAGCAGTTTAAACGAGAGAGCTGGCAATGGAATTGGAGAGAGCATCATGTTCAGAGAAC 1920
DB 621 LysGlnPheLysArgArgSerCySGlnMetGlnPheGlyGlnSerIleMetSerGlnLysn 640
QY 1921 AGGTCAACGAGAGAGCTGGCGAAAAGTGGCGAGTCACTGACTTTTCGGCGAGCATGGAA 1980
DB 641 ArgSerArgGlnGlnLysGlnValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCATTTAGAGTCTCC 1995
DB 661 IleIleGlnValSer 665

RESULT 4

AAU09016

ID AAU09016 standard; protein; 665 AA.

AC AAU09016;

DT 18-DEC-2001 (first entry)

DE Human dual specificity phosphatase 21117.

KW Human, dual specificity phosphatase 21117; hepatotropic; cytosolic;
KW hematopoietic disorder; autoimmune disorder; diabetes mellitus;
KW rheumatoid arthritis; multiple sclerosis; Crohn's disease;
KW cellular proliferative; differentiative disorder; haemolytic anaemia;
KW acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer;
KW immunogen.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Domain 11..131

FT Domain 158..297

FT Active-site 242..254

FT /label= Tyrosine_specific_protein_phosphatase_active_site

PN W0200173059-A2.

PD 04-OCT-2001.

PF 23-MAR-2001; 2001WO-US009477.

PR 24-MAR-2000; 2000US-0191858P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Meyers RA;

DR WPI; 2001-611635/70.

DR N-PSDB; AAS14639.

PT New human dual specificity polypeptides and nucleic acids for diagnosis

PT of disease and treatment of e.g. liver disorders.

PS Claim 9; Fig 1; 143pp; English.

CC The invention relates to two novel human dual specificity phosphatases
CC designated 21117 and 38692, the nucleic acids encoding them (including
CC fragments, allelic variants, their complements or nucleic acids that
CC hybridise to them) and antibodies raised against the proteins. The

CC antibody is useful for detecting the presence of the polypeptide, and the
CC nucleic acid fragments are useful for detecting the presence of the
CC nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
CC antisense sequences) are useful for modulating the activity or expression
CC of the polypeptide/nucleic acid, useful for the treatment of e.g.
CC hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
CC listed in the specification) liver disorders, erythroid associated
CC disorders (e.g. haemolytic anaemia) cellular proliferative or
CC differentiative disorders, leukemias (e.g. acute myeloid leukaemia),
CC carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and
CC 38692 are also useful for modulating the proliferation, survival,
CC migration or differentiation of a 38692 or 2117-expressing cell. The
CC polypeptide and nucleic acids are useful for identifying modulating
CC agents. The present sequence represents the dual specificity phosphatase
CC 21117

SQ Sequence 665 AA;

Alignment Scores:

Pred. No.:	3,16e-287	Length:	665
Score:	3406.00	Matches:	663
Percent Similarity:	99.85%	Conservative:	1
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	94.22%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345A-108_COPY_538_2532 (1-1995) x AAU09016 (1-665)

QY 1 ATGGCCATGAGATGATTTGAAATGAAATTTGTAAGAGTGGTGGCTGGAGAA 60
DB 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuGln 20
QY 61 AGTGAACGAGAAAAGTGTCTGTAATGATGATCCGCGCATTTGTGGAATCAATACATCC 120
DB 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnTyrAsnThrSer 40
QY 121 CACATTTTGAAGACCATTAATATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACG 180
DB 41 HisIleLeuGlnAlaIleAsnIleAsnCySerIleLeuMetLysArgArgLeuGln 60
QY 181 GACAAAGTGTAAATTAAGAGCTCATCCAGCATTTGACGAAACATTAAGTTGAT 240
DB 61 AspLysValIleuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY 241 TGCAGTCAAGAGCTGTGATTTACATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 300
DB 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTCTCACTGACTTCTGGGTAAACTGAGAGAGAGCTTCACTCTGTTACCTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGTGGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 121 LeuAlaGlyIlePheAlaGlnPheSerArgCysPheProGlyLeuCySGlnGlyLysSer 140
QY 421 ACTTAGTCCCTACCTGATTTCTCAGCCTTGCTTACCTGTTGGCAACATTTGGCCAAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAACTTTATCTTGGCTGCGCAGAGATGCTCCCAACAGAGAGATGA 540
DB 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgPheValLeuAsnLysGlnLysLeuMet 180
QY 541 CAGCAGAAATGGAGTTGTTATGTGTAAATGCGCAGCATATCTGTCCAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCAATTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 220
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu

```

QY 661 CCGTGGTGGACAAATCAGTAGATTTCATTTAGAGAAAGCAAAAGCCTCAATGGATGTT 720
    |||||
Db 221 ProTPrneuapbyservValaephneillegluValaValaaservnglyCysVal 240
QY 721 CTAGTCACTGTTTAACTGGGAATCTCCCGCTCCGCCACATCGCTATCGCTCATCATG 780
    |||||
Db 241 LeuValHieCysLeuValaglyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIleMet 260
QY 781 AAGAGATGAGACATGTCCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAAGACTACT 840
    |||||
Db 261 LybArgMetCaspMetSerLeuapbgluValArgPheVallybgluValysArgProThr 280
QY 841 ATATCTCCAAACTTCAATTTCTGGGCCAATCTCTGAGATGATGAGAAAGATTAAGAAC 900
    |||||
Db 281 IleSerProapbneapbneapbneapbneapbneapbneapbneapbneapbneapbne 300
QY 901 CAGACTGAGACATCAGGGCCAAAGACAACTCAAGCTGCTGACCTGAGAGAACCAAT 960
    |||||
Db 301 GluThrGlyAlaSerGlyProlybSerLybLeuLybLeuHieIleuGlyIleProapb 320
QY 961 GAACCTGCTCTGCTGCTCTAGAGGGGTGACAGAAAAAGCAGACCCCTCATCTCAACC 1020
    |||||
Db 321 GluProValProAlaValSerGlyGlyGlyIleSerGlyIleThrProLeuSerProPro 340
QY 1021 TGTGCGGACTGCTGCTGCTGCTGAGAGGACGAGCAAAAGGCGCGCTCCGCGAGGCTG 1080
    |||||
Db 341 CysAlaIlePserAlaThrSerGlyAlaIleGlyIleThrProValIleSProAlaSerVal 360
QY 1081 CCCAGCGTGGCCAGCGCTGACGCGCTGCTGTTAGAGACAGCCCGCTGTATACAGGCGCTC 1140
    |||||
Db 361 ProSerValProSerValGlnProSerLeuLeuGluIleapSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGGCTGACCTGCTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCTTCTCTTC 1200
    |||||
Db 381 SerGlyLeuHieIleuSerAlaIleapArgLeuGluIleapSerAlaIleuLybArgSerPhe 400
QY 1201 TCTCTGATATCAATCAGTTTATATTCAGGAGATGACAGATGACATCTTACATGGCTTC 1260
    |||||
Db 401 SerLeuapbIleLybSerValIleSerIleSerAlaSerMetAlaIleSerLeuHieIleGlyPhe 420
QY 1261 TCCTCATCAGAAAGATGCTTGAATACAACTTCCATCTGATGAGGAGCAAC 1320
    |||||
Db 421 SerSerSerGluIleapbAlaLeuGlyIleThrLybProSerThrThrLeuapbGlyThrIleapb 440
QY 1321 AAGCTATGCCAGTTCCTCCCTGTTCAAGAACTATCGAGAGACACTCCCGAAACAGTCTCT 1380
    |||||
Db 441 LybLeuCybGlnIlePheSerProValGlnGlyIleuSerGlyIleThrProGlyIleThrSerPro 460
QY 1381 GATTAAGAGAGAGACGACATCCCAAGAGGCTGACAGCGCGCGCTTACAGACGACG 1440
    |||||
Db 461 ApbLybGlyGluIleapbIleProLybLybLeuGlnIleThrAlaArgProSerApbSerGln 480
QY 1441 AGCAAGCAGATTGATTCGCTCAGAACACGACGACGATGCGACCGCCAGAGATCCCTTTTA 1500
    |||||
Db 481 SerLybArgLeuHieIleSerValArgThrSerSerSerGlyThrAlaGlnIleapbSerLeuLeu 500
QY 1501 TCTCCATGTCATGAAGTGGAGCGGTGAGAGACAAATTAACAACAGCTTCTTTTCGGC 1560
    |||||
Db 501 SerProLeuHieIleapbArgSerGlySerValGluIleapbAlaIleThrIleThrSerPheLeuPheGly 520
QY 1561 CTTTCACACGACGACGACGACCTCAGAGATGCTGCGGCGCTGAGGCTTAAAGGCTGGCAC 1620
    |||||
Db 521 LeuSerThrSerGlnIleGlnIleuThrLybSerAlaGlyIleuGlyIleuLybGlyIlePheIle 540
QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACGACGAGCTGTATTTGGC 1680
    |||||
Db 541 SerAlaPheIleuValProGlnIleThrThrProSerLeuThrSerSerThrPheAla 560
QY 1681 ACGAGTCTCACAACATTTACTGCTGCTCAGCCATCTACGAGAGGACGATGCTACTCT 1740
    |||||
Db 561 ThrGlySerSerIlePheThrIlePheThrIlePheThrIlePheThrIlePheThrIlePheThr 580
QY 1741 GCCTACAGCTGACGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800

```

```

Db 581 AlaTySerCysSerGlnLeuProThrCybGlyApbGlnValTySerSerValArgArg 600
QY 1801 CAGAACCAAGTACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGACCCCTTTGAA 1860
    |||||
Db 601 GlnLybProSerApbArgAlaApbSerArgArgSerThrPheIleGlyIleuSerProPheGly 620
QY 1861 AAGCACTTTAAACGAGAGCTGCGCAATGGAATTTGGAGAGACATCATGTCAGAGAAC 1920
    |||||
Db 621 LybGlnPheLybArgArgSerCybGlnMetGluPheGlyGluSerIleMetSerGluIle 640
QY 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGCAGTCACTTACCTTTTCGGGACAGAGAA 1980
    |||||
Db 641 ArgSerArgGluGluLeuGlyIleValGlySerGlnSerSerPheSerGlySerMetGlu 660
QY 1981 ATCATGAGCTCTCC 1995
    |||||
Db 661 IleIleGluValSer 665
    |||||

RESULT 5
AAU79156
ID AAU79156 standard; protein; 665 AA.
XX
AC AAU79156;
XX
DT 02-JUL-2002 (first entry)
XX
DE Human dual-specificity phosphatase-3 (DSP-16) protein.
XX
KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW cancer; graft-versus-host disease; allergy; metabolic disease;
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
KW intercellular adhesion; DSP-16 modulator; chromosome 12p.
XX
OS Homo sapiens.
XX
FH
FT Key Location/Qualifiers
FT Domain 242..251
FT /label= Active_site_domain

MO200226997-A2.
PD 04-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US030124.
XX
PR 26-SEP-2000; 2000US-0235487P.
XX
PA (CEPT-) CEPTYR INC.
XX
PI Luche RM, Wei B;
XX
DR MPI: 2002-315802/35.
XX
DR N-PSDB; ABK47596.
XX
PT New DSP-16 polypeptide, useful for identifying modulators of its
PT activity, which can be used in the treatment of disorders such as
PT Duchenne muscular dystrophy, or cancer.
XX
PS Claim 1; Fig 2; 87pp; English.
XX
CC The present invention relates to a new polypeptide, DSP-16, having a 665
CC amino acid sequence, given in the specification, or a variant having at
CC least 50 % identical residues, which retains the ability to
CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
CC activity can be used for identifying agents which modulate DSP-16
CC activity, for modulation of a proliferative response in a cell, survival
CC of a cell, or differentiation of a cell. The cell displays contact
CC inhibition of cell growth or anchorage independent growth and may display
CC altered intercellular adhesion. The agent may modulate apoptosis, or the
CC cell cycle. The identified modulators can be used to treat Duchenne

```


QY	1921	AGGTCACGGGAAGAGCTGGGGAAGATGGGACGTACGTTACTTTTCGGGACAGCATGGA	1980
DB	641	ArgSerArgGluGluLeuGluLysValGlySerGlnSerSerPheSerGlySerMetGlu	660
QY	1981	ATCATTTGAGCTCTCC	1995
DB	661	IleIleGluValSer	665
RESULT 6			
AAU09946			
ID	AAU09946	standard; protein, 665 AA.	
XX	AAU09946;		
AC			
XX			
DT	18-JUN-2002	(first entry)	
XX			
DE	Protein sequence of human (dual specificity phosphatase) DUSP-10.		
XX			
KW	Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke;		
KW	neuronal degeneration syndrome; Alzheimer's disease; depression;		
KW	schizophrenia; asthma; immune disorder; inflammatory process; arthritis;		
KW	osteoporosis; diabetes.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200177340-A1.		
XX			
PD	18-OCT-2001.		
XX			
PF	06-APR-2001; 2001WO-EP003966.		
XX			
PR	10-APR-2000; 2000EP-00107143.		
XX			
PA	(WERE) MERCK PATENT GMBH.		
PI			
XX	Duecker K;		
DR	WPI; 2002-010917/01.		
XX			
DR	N-PSDB; AAS15768.		
XX			
PT	Novel dual specificity phosphatase polypeptides useful for treating		
PT	cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's		
PT	disease, depression, schizophrenia, asthma and immune disorders.		
XX			
PS	Claim 2; Page 37-39; 43pp; English.		
XX			
CC	The present invention relates to a new isolated dual specificity		
CC	phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid		
CC	sequence that is fully defined in the specification. The invention also		
CC	provides a sequence encoded by a 309 nucleotide sequence fully defined		
CC	in the specification, and a sequence having at least 95 % identity to the		
CC	polypeptide, or fragments or variants of DUSP-10. The invention is useful		
CC	for treating cancer e.g. leukemia, colon carcinoma, lung cancer,		
CC	prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,		
CC	stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,		
CC	schizophrenia, cardiac myopathies, asthma, immune disorders,		
CC	inflammatory processes e.g. arthritis, bowel disease, type I diabetes,		
CC	osteoporosis, diabetes and diabetes associated diseases. The molecules of		
CC	the invention are also useful as vaccines for inducing immunological		
CC	response in a mammal, in disease diagnosis and in assays for screening		
CC	agonistic or antagonistic compounds. Other uses of the invention include		
CC	identifying membrane bound or soluble receptors, as a diagnostic reagent,		
CC	in chromosome localisation studies, and as a valuable tool in tissue		
CC	expression studies. The present sequence represents the dual specificity		
CC	phosphatase, DUSP-10, protein of the invention		
XX			
XX	Sequence 665 AA;		
XX			

Alignment Scores:	
Pred. No.:	3,166-287
Score:	3406.00
Percent Similarity:	99.85%
	Length: 665
	Matches: 663
	Conservative: 1

Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	94.22%	Indels:	0
DB:	5	Gaps:	0

5)

Qy	1	ATGGCCCATGAGATGATTTGAAAGCTCAAAATTGTACTGAGAGGTTGGAGCTCTGCTGGAA	60
Db	1	MetAlaHsGluMetC1LeG1YThrGln1LeVal1ThrGluArgLeuValAlaLeuLeuGlu	20
Qy	61	AGTGAACGGGAAAAAGTGTGCTGCTAAATTGATAGCCGGCCATTGTGTGAATACAAATACATCC	120
Db	21	SerC1YhrGluYsValLeuLeuLe1LeaPserArgProPheValGluTYrAsnThsSer	40
Qy	121	CACATTTTGGAAAGCCATTATATCACTGCTCCAGCTTATGAAAGCGAAGTTGCACAG	180
Db	41	HisLeuGluAla1Lea1LeaMet1LeaAsnCysSerYsLeuMetC1YsArgLeuGlnGln	60
Qy	181	GACAAABGTGTTAATTAGAGGCTCATCAGAGATTCAAGGAAACATAAGCTTGACATTGAT	240
Db	61	AspYsValLeuLeu1LeuTrGluLeuLe1LeGlnHisSerAlaYsHisYsValAsp1LeaSP	80
Qy	241	TGCAGTCAGAAAGTTGTAGTTTACATCAAAAGCTCCCAAGATGTGTGCTCTCTCTCA	300
Db	81	CysSerGlnYsValValValTYrAspGlnSerSerGlnAspValAlaSerLeuSerSer	100
Qy	301	GACTGTTTTCTCACTGTACTTCTGGGTAACTGGAGAAAGAGCTTCAACTCTGTTCACTG	360
Db	101	AspYsPheLeuThrValLeuLeuG1YysLeuGlnYsSerPheAsnSerValHisLeu	120
Qy	361	CTTGCAGGTGGGTTGTCGAGGTTCTCGTGGTTTCCCTGGCCCTGTGAAAGGAAATCC	420
Db	121	LeuAlaG1YpPheAlaGluPheSerArgCysPheProG1YLeuCYsGluG1YysSer	140
Qy	421	ACTCTAGTCCCTCACTCATTTCTCAGCCTTGCTTACTCTGTGTGCAACATTTGGGCCAAC	480
Db	141	ThrLeuValProThrCYsIleSerGlnProCYsLeuProValAlaLeuMet1LeG1YProThr	160
Qy	481	CGAATTTCTTCCCAATCTTTATCTTGGGCTGCCAGAGAGATGTCCTCAACAGAGACTGATA	540
Db	161	Arg1LeuProAsnLeuTYrLeuG1YCYsGlnArgAspValLeuAsn1YsGluLeuMet	180
Qy	541	CAGAGAAATGGAGATTTGGTTATGTGTTAAATGCCAGCTATACCTGTCCAAAGCTCGACTTT	600
Db	181	GlnGlnAsnG1Y1LeG1YTYrValLeuAsnHisAsenThrCYsProYsPheAspPhe	200
Qy	601	ATCCCCAGTCTCATTTCTCGCGTGCTCCTGTGATGACAGCTTTTGTGAGAAATTTTG	660
Db	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCYsGluTYrIleu	220
Qy	661	CCGTGTTTGGACAAATCACTAGATTTTCAATTGAGAAAGCAAAAGCTCCATAGATGTCT	720
Db	221	ProTrPheAsnAspYsSerValAspPhe1LeG1YysAlaYsAlaSerAsnG1YCYsVal	240
Qy	721	CTAATGACAGTTTAGCTGGGATCTCCCGCTCCGCCACATCGGTATCGCTCAATCATG	780
Db	241	LeuValHisCYsLeuAlaG1Y1LeSerArgSerAlaThr1LeaAla1LeaTYrIleMet	260
Qy	781	AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAAGACTACT	840
Db	261	YsArgMetAspMetSerLeuAspGluAlaTYrArgPheValYsGluYsArgProThr	280
Qy	841	ATATCTCCAAACTTCATTTTCTGGGCCAACTCCTGGACTRTAGGAAGAATTTAAGAAC	900
Db	281	IleSerProAsnPheAsnPheLeuG1YGlnLeuLeuAspTYrGluYsYs1LeYAsn	300
Qy	901	CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGAGAGACCAAT	960
Db	301	GlnThrC1YAlaSerG1YProYsSerTYsLeuYsLeuMetHisLeuGluYsProAsn	320
Qy	961	GAACCTGTCCCTGTGCTCTCAGAGGGTGAACAGAAAAGAGAGCGCCCTCAGTCCACC	1020
Db	321	GlnProValProAlaValSerGluG1YG1YGlnYsSerGluThrProLeuSerProPro	340

QY	1021	TCGCGCGACTTCGCTACTTCAGAGCGCAGAGCAAAAGGCGCGTGCATCCGCGCAGCGTG	1080
Db	341	CysAlaAspSerLeuAlaThrSerGluAlaIaGluInaArgProValHisProAlaSerVal	360
QY	1081	CCCAAGCGTCCCAAGCGCTTGAGCGCGTGGCTTTAGAGACAGCCGCTGGTACAGCGCTC	1140
Db	361	ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu	380
QY	1141	AGNGGCGTGCACCTGTCCGCGACAGCGCTGGAAAGACAGCAATAAGCTCAAGCGTTCCCTTC	1200
Db	381	SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerLeuGluValGlnAlaLeu	400
QY	1201	TCTCTGATATCAATCAGTTTCATATTCACCGAGCATGGGACGATCTTACATGGCTTC	1260
Db	401	SerLeuAspIleLeuSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe	420
QY	1261	TCCCATAGAGAAGTGCCTTTGGAAATCTAACAACTTCCACTCTCTGATGGGAGCAAC	1320
Db	421	SerSerSerGluAspAlaLeuGluTyrTyrLeuProSerThrThrLeuAspGlyIlnAsn	440
QY	1321	AAGCTATGCCAGTTCTCTCCCTGTTCAGAGACTATTCGAGAGACTCCGAAACCAAGTCT	1380
Db	441	LysLeuLysGlnPheSerProValGlnGluLeuSerGlnThrProGluThrSerPro	460
QY	1381	GATTAAGAGAGAGACCGACATCCCGAAGAAAGCTGCAGACCCGACGCTTCAGACGCCAG	1440
Db	461	AspLysGlnGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln	480
QY	1441	AGCAACCGATTGCATTCGGTGCAGAAACAGACACAGCGGACCGCCAGACGTCCTTTTA	1500
Db	481	SerLysArgLeuHisSerValArgThrSerSerGlyIlnAlaGlnArgSerLeuLeu	500
QY	1501	TCTCCACTGCATCGAAAGTGGAGCGGTGAGAGCAATTACACACCGCTTCTTTGCGC	1560
Db	501	SerProLeuHisArgSerClySerValGlnAspAsnTyrHisThrSerPheLeuPheGly	520
QY	1561	CTTTCCACACGCGCAGACGACCTCAAGAACTGTGCTGGCGCTGAGCGCTTAAGCGCTGGCAC	1620
Db	521	LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheHis	540
QY	1621	TCGGATATCTTGCGCCCCCAGACCTTACCCCTTCCCTGACACAGACCTGGTATTTTGCGC	1680
Db	541	SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrIyrPheAla	560
QY	1681	ACAAGTCCACACTTCTACTGTGCTGCCTCAAGCATTCACGAGGAGCGAGTGCAGTTACTCT	1740
Db	561	ThrIleuSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSer	580
QY	1741	GCTTACAGCTGCAGCCAGCTGCCACATCTTGCGAGACCAAGTCTATTCTTGCGCAGCGCG	1800
Db	581	AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg	600
QY	1801	CAAGAAGCAAGTACAGAGACTGACTCGCGCGGAGCTGCATGAGAGAGAGAGAGAGAGAGAG	1860
Db	601	GlnLysProSerAspArgAlaAspSerArgArgSerTyrPheGlnGluSerProPheGln	620
QY	1861	AAGCACTTTAAACGAGAAAGCTGCCAAATGTGAATTGGAGAGAGCATATCTACAGAAAC	1920
Db	621	LysGlnPheLysArgSerCysGlnMetGlnPheGlyGlnSerTyrIleMetSerGlnAsn	640
QY	1921	AGGTACACGGGAAGACTCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTGGGCGAGATGAA	1980
Db	641	ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln	660
QY	1981	ATCATTTAGGTCTCC	1995
Db	661	IleIleGlnValSer	665

AC	AAU575789;	
AD	08-MAY-2002 (first entry)	
DE	Human protein phosphatase 7 (PP7) protein sequence.	
XX		
XX		
XX	Human, protein phosphatase 7; PP7; immune system disorder; AIDS;	
KW	acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma	
KW	Crohn's disease; neurological disorder; epilepsy; Huntington's disease	
KW	dementia; Parkinson's disease; developmental disorder; Down's syndrome	
KW	cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymph	
XX	melanoma; myeloma sarcoma.	
OS	Homo sapiens.	
XX		
XX		
Key	Location/Qualifiers	
FH	Domain	11..131
FT	/label="Rhodanese_like_domain	15..170
FT	/note="Dual specificity protein phosphatase"	85..298
FT	Region	158..297
FT	/note="VH1-type dual specificity protein phosphatase"	158..297
FT	/label="Catalytic domain DSPC	220..281
FT	/note="Dual specificity phosphatase"	220..281
FT	/note="Tyrosine specific protein phosphatases active	242..254
FT	Region	242..254
FT	/note="Tyrosine phosphatase"	242..252
FT	/note="Tyrosine specific protein phosphatases"	
PN	MO200210363.A2.	
XX		
XX	07-FEB-2002.	
XX	26-JUL-2001; 2001WO-US023716.	
XX	28-JUL-2000; 2000US-0221619P.	
XX	03-AUG-2000; 2000US-0223372P.	
PR	10-AUG-2000; 2000US-0224309P.	
PR	18-AUG-2000; 2000US-0226728P.	
PR	30-AUG-2000; 2000US-0229254P.	
PR	08-SEP-2000; 2000US-0231366P.	
XX		
XX	(INCY-) INCYTE GENOMICS INC.	
PI	Tang YF, Elliott VS, Rankumar J, Yao MG, Burford N, Wang YE;	
PI	Stewart EA, Gandhi AR, Patterson C, Lee EA, Hatalla AJA, Lu DM;	
PI	Trinbouley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;	
XX	Walia NK, Kearney L;	
XX	MP1; 2002-188735/24.	
DR	N-P5DB; ABK14474.	
XX		
PT	New protein phosphatases, useful for diagnosing, treating or preventing	
PT	immune system disorders (e.g. Crohn's disease), neurological disorders	
PT	(e.g. Parkinson's disease), or cell proliferative disorders (e.g.	
PT	cancers).	
XX		
PS	Claim 1; Page 107-108; 117pp; English.	
XX		
CC	The present invention relates to a new polypeptide, a naturally occurring	
CC	amino acid sequence at least 95 % identical to it, a biologically active	
CC	fragment of it or an immunogenic fragment of it, the polypeptides,	
CC	polynucleotides, agonists and antagonists are useful for diagnosing,	
CC	treating or preventing disorders associated with aberrant expression of	
CC	protein phosphatases (PP), particularly immune system disorders e.g.	
CC	acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,	
CC	asthma or Crohn's disease, neurological disorders e.g. epilepsy,	
CC	Huntington's disease, dementia or Parkinson's disease, developmental	
CC	disorders e.g. Down's syndrome, or cell proliferative disorders e.g.	

CC cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma
CC or sarcoma. The present amino acid sequence represents human protein
CC phosphatase 7 (PP7) which is one of several human protein phosphatases
CC (AAU75783-AAU75792) of the invention

XX Sequence 665 AA:

Alignment Scores:

Prod. No.:	3.16e-287	Length:	665
Score:	3406.00	Matches:	663
Percent Similarity:	99.85%	Conservative:	1
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	94.22%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-108_COPY_538_2532 (1-1995) X AAU75789 (1-665)

```
QY 1 ATGGCCATGAGATGATGGAGCTCAATTTGTTACTGAGAGGTGGTGGCTGTGGAA 60
DB 1 MetAlaHisGluMetCileGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAGAGCTGCTTAATTGATAGCCGACATTTGTGCAATACATACC 120
DB 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
QY 121 CACATTTGGAGCGCATTAATATCACTGCTCCAGCTTATGAAGGAGTTGCAACAG 180
DB 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerIleValMetCysArgGluGln 60
QY 181 GACAAAGTTTAATTCAGAGCTATCCAGCATTCAGCGAAGCAATAGTTGACATTGAT 240
DB 61 AspIleValLeuIleThrGluLeuIleGlnHisSerIleValHisValAspIleAsp 80
QY 241 TCGAGTCAGAAAGTTGATGTTACGATCAAAAGCTCCCAAGATGTGCTCTCTTCA 300
DB 81 CysSerGlnIleValValValIleValAspGlnSerSerGlnAspValAlaSerLeuSer 100
QY 301 GACTGTTTCTCACTGATCTTGGGGTAACTGAGAGAGCTTCACTGTTCACCTG 360
DB 101 AspCysPheLeuThrValIleLeuGluGlyLysLeuGlnIleValSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
DB 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProIleLeuGlyGlyLysSer 140
QY 421 ACTTGAAGTCCCTACCTGATTTTCAAGCTTGTGCTTACCTTGTGCAATTTGGCCAA 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAAATCTTCCCAATCTTTATCTTGGCTGCGACAGAGATGCTCCCAAGAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuValIleGluGlyCysGlnArgAspValLeuAsnIleGluMet 180
QY 541 CAGCAGAAATGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
DB 181 GlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerAsnThrCysProIleProAspPhe 200
QY 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAAATGACAGCTTTTGTGAAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
QY 661 CGGTGGTGGAGCAATCAGATGATTTTCATGAGAAAGCAAAAGCTCCCAATGATGTT 720
DB 221 ProIlePheAspIleValAspPheIleGlnIleValAlaValAsnArgGlyCysVal 240
QY 721 CTAGTCACTGTTAGTGGATGATCCCGCTCCGACCATCGCTATCGCTACATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIle 260
QY 781 AAGAGATGACATGTTGTTAGATGAAGCTTACAGATTTGTGAAAAGAAAGACCTACT 840
DB 261 LysArgMetAspMetSerLeuAspGlnAlaIleValIleValIleValIleValIleValIle 280
```

```
QY 841 ATATCTCAAACTTCAATTTTCTGGGCCAACTCTGTGATCATGAGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIleValIleValIleValIle 300
QY 901 CAGACTGAGCATCAGGGCCCAAGAGCAAACTCAAGCTGTGCACTTGAGAAAGCCAAAT 960
DB 301 GlnThrGlyAlaSerGlyProIleValSerIleValLeuValLeuValIleValIleValIle 320
QY 961 GAACCTGTCCCTCTCTCTCAAGAGGTGAGCAAGAAAGAGAGAGCCCTCACTCCACC 1020
DB 321 GlnProValProAlaValSerGlnIleGlyGlnIleValSerGlnIleValSerProPro 340
QY 1021 TGTGCGACCTGTCTCTCTCAAGAGGAGAGAGCAAGAGCCCGCTGTGATCCCGCAG 1080
DB 341 CysAlaAspSerAlaThrSerIleValAlaIleGlyGlnArgProValHisProAlaSerVal 360
QY 1081 CCCAGCGTCCAGCGTCCAGCGCTGTGTTAGAGAGAGCCCGCTGTGATCAGCGCTC 1140
DB 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGCTGACACTGTCTCCCGAGCAGAGCTGGAAGAGAGCAATAGCTCAAGCGTCTTC 1200
DB 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnIleValIleValIleVal 400
QY 1201 TCTCGATATATCAATCAGTTTCAATTCAGCAGATGGACATCTTACATGCGCTTC 1260
DB 401 SerLeuAspIleValSerValIleValIleValIleValIleValIleValIleValIle 420
QY 1261 TCTCATCAGAAAGATCTTTGGAATCTAACAACTTCCATCTGTGATGGAGCAAC 1320
DB 421 SerSerSerGlnAspAlaLeuGlnIleValIleValIleValIleValIleValIleVal 440
QY 1321 AAGCTATGCCAGTCTTCCCTGTTCAGAGAACTTCGAGAGAGACTCCCGAAACAGATCCT 1380
DB 441 LysLeuCysGlnPheSerProValGlnGlnLeuSerGlnIleThrProGlnIleThrSerPro 460
QY 1381 GATPAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 461 AspIleGlnIleValIleValIleValIleValIleValIleValIleValIleValIle 480
QY 1441 AGCAGAGATGATTCGATTCGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 481 SerIleArgLeuHisSerValIleValIleValIleValIleValIleValIleValIle 500
QY 1501 TCTCCATGCTATGAGAGTGGAGAGCTGAGAGAGCAATTAACACAGCTTCTTTCGGC 1560
DB 501 SerProLeuHisArgSerGlySerValGlnAspAsnIleThrIleSerPheLeuPheGly 520
QY 1561 CTTCACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB 521 LeuSerThrSerGlnIleHisLeuIleValIleValIleValIleValIleValIleVal 540
QY 1621 TCGGATATCTTGGCCCGCCAGAGAGCTTACCCCTTCCCTGACAGAGAGCTGGTATTTTGGC 1680
DB 541 SerAspIleLeuAlaProGlnIleThrIleThrIleThrIleThrIleThrIleThrIle 560
QY 1681 ACAGAGTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCT 1740
DB 561 ThrGlnSerSerHisPheIleValIleValIleValIleValIleValIleValIleVal 580
QY 1741 GCCTACAGCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
DB 581 AlaIleValSerCysSerGlnIleProIleValIleValIleValIleValIleValIleVal 600
QY 1801 CAGAAAGCAAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
DB 601 GlnIleValProSerAspArgAlaAspSerArgArgSerIleIleGlnIleValSerProPheGln 620
QY 1861 AAGCAGTTTAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
DB 621 LysGlnPheLysArgIleValIleValIleValIleValIleValIleValIleValIle 640
QY 1921 AGGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
```

D6
641 ArgertergclutluenuengllyrsValGlsySerInserSerPhseRcglYsrMerCln 660

OY 1981 ATCAATTAGGTCTCC 1995
D6 661 lletleclulvalser 665

RESULT 8
ABB97946
ID ABB97946 standard; protein; 665 AA.
AC ABB97946;
XX
XX
DT 06-SEP-2002 (first entry)
DE Human protein sequence #13.
KW Human; brain; tonsil; hippocampus; foetal brain; diagnosis.
OS Homo sapiens.
PN WO200252005-A1.
PD 04-JUL-2002.
PF 20-DEC-2001; 2001WO-JP011217.
PR 22-DEC-2000; 2000JP-00389742.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
PI (CELFE-) CELESTAR LEXICO-SCI LTD.
PI Ohara O, Nagase T, Nakajima D;
WP1; 2002-500762/53.
DR N-PSDB; ABN83966.

Genes and their expression products cloned from human cDNA libraries for treatment and diagnosis of diseases associated with their expression.
Claim 1(a); Page 112-116; 238pp; Japanese.

The invention relates to DNA encoding polypeptides directly cloned from CC cDNA libraries originating in adult whole brain, human tonsil, human CC adult hippocampus and human foetal whole brain. Polypeptides and CC polynucleotides of the invention may be used in the investigation of CC differential expression of the DNA sequences in normal subjects and CC oligonucleotide probes and DNA chips for diagnosis and identification of CC drugs for treatment of diseases with which the DNA sequences are CC associated. The sequences given in records ABB97934-ABB97964 represent XX

SQ Sequence 665 AA;

Alignment Scores:
Pred. No.: 3.16e-287
Score: 3406.00 Length: 665
Percent Similarity: 99.85% Matches: 663
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 94.22% Indels: 1
DB: 5 Gaps: 0

US-10-029-345A-108_COPY_538_2532 (1-1995) x ABB97946 (1-665)

OY 1 ATGGCCCATGATGATTTGAACATAATTTTACTGAGAGTTGGCGCTTCCTGGAA 60
D6 1 MetAlAHtslglwettllleglThrglnlleValnrlrGlualrgLeuValAlaleuLeugln 20
OY 61 AGTGAACCGGAAAAAGtGCTGCTTAATTGATAGCCGGCCATTGTGTAATCAATACATCC 120
D6 21 SerCljThrcIulysValneuLeuileaspetrArPrropheValGIuTYrrAntHrSsr 40

Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 94.22%
DB: 5

Conservative: 1
Matches: 1
Indels: 0
Gaps: 0

US-10-029-345a-108_COPY_538_2532 (1-1995) x AAU79929 (1-665)

QY 1 ATGGCCATGAGATGGAACCAATTGTTACTAGAGGTTGGCTCTGCTGGA 60
DB 1 MetalaHsglumetIIeglyThnglnllevalThrgluarGheValAlaleuengl 20
QY 61 AGTGAACGGAAAAGCTGCTTAATTGATAGCCGGCCATTGTGGAATACATCATCC 120
DB 21 SerGlyThngluysValleuLeuIIeasPserArGProheValGluYrasmhser 40
QY 121 CACATTTTGAAGCATTAAATCACTGCTCCAGCTTATAGAGCAAGGTTGCACAG 180
DB 41 HisleleuenglualleleamIleamCysSerLysleuMetLysArgArgleuGln 60
QY 181 GACAAAGCTTAATTACAGACTCATCCAGATTGACGAAACATPAGATTGACATTGAT 240
DB 61 AspLysValleuIIeThrgluLeuIIeglnHsserAlalyshleLysValAspIleasP 80
QY 241 TGCAGTCAGAAAGTTTGTAGATTACCATCAAGCTCCCAAGTTGCTCTCTCTCA 300
DB 81 CysSerGlnLysValAlValAlYrAspGlnSerSerGlnAspValAlAserLeuSerSer 100
QY 301 GACTGTTTCTGACGTACTTCTGGGTTAACTGAGAAAGACTTCAACTCTGTTCACTG 360
DB 101 AspCysPheleuThValleuLeuIdlyLysleuGluLysSerPheamserValhshleu 120
QY 361 CTTCAGAGTGGTTTGTGAGTCTCTGTTGTTTCCCTGCTCTGTGAAGAAATTC 420
DB 121 leuAlaglyGlyPheAlagluPheSerArgCysPheProGlyLeuGlyLysSer 140
QY 421 ACTTAGGCTTACCTGATTTCTGACCTTGTCTTACCTGTTGCCAAATGGGCAACC 480
DB 141 ThreleuValProhmCysIleSerGlnProCysleuProValAlAsnIleGlyProhm 160
QY 481 CGAATTTCTTCCCAATCTTATCTGCTGCTGACGAGATGCTCTCAACAGAGCTGTAT 540
DB 161 ArgIleleuProleuLeuYrleuGlyCysGlnAspValleuAsnlyLysleuMet 180
QY 541 CAGCAGATGGGATGGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCTGATCTT 600
DB 181 GlnIlnAsnIlyIleGlyYrValleuAsnAlAserAnthrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 201 IleProGlnSerHisPheleuArgValProValAsmAspSerPheCysGluYrIleleu 220
QY 661 CCGGTTTGGGCAAAATCATGATTTTATTGAGAAAGCAAAAGCTCCCATGATGATGTT 720
DB 221 ProTrpPheAspLysSerValAspPheIleGluYrAlalySalAserAnthrCysVal 240
QY 721 CTATGCTACTGTTTATGCTGGATCTCCGCTCCGCAACATCCGTATCCCTTCAATCAT 780
DB 241 leuValhshCysleuAlaglyYrIleSerArgSerAlathrIleAlleAlalyrIleMet 260
QY 781 AAGAGATGAGCATGCTTTTATGATGAGCTTACGATTTGTGAAGAAAGCAAAAGCTTACT 840
DB 261 LysArgMetAspMetSerleuAspGluAlalyrArgPheValLysGluYrAspArgProhm 280
QY 841 ATATCTCCAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 281 IleSerProleuPheAspPheleuGlyGlnleuLeuAspYrGluYrLysIleLysAsn 300
QY 901 CAGACTGAGCATAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTGGAAGAGCAAT 960
DB 301 GlnhThrglyAlaserglyProLysSerLysleuYrleuLeuHshleuGluYrProhm 320
QY 961 GAACCTGCTCTGCTGCTGCTGAGAGGTGAGCAAGAAAGGAGACCCCTCACTCAACC 1020

DB 321 GluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProPro 340
QY 1021 TGGCCAGCTCTGTACTTACCTCAGAGGACAGGACAAAGCCCGTGCATCCCGCAGCGTG 1080
DB 341 CysAlAspSerSerAlathrSerIlnAlalaglyGlnArgProValhshProAlAserVal 360
QY 1081 CCCAGCTGCCAGGCTGACGCCGTGCTGTTTATAGAGCAAGCCCGCTGTTACAGGCGCTC 1140
DB 361 ProSerValProSerValGlnProSerleuLeuIlnAspSerProleuValGlnAlaleu 380
QY 1141 ACTGGGCTGCACTGTGCTGGCAGACAGCTGGAAGACAGCAATAAGCTCAAGCTTCTTC 1200
DB 381 SerGlyleuHshleuSerAlAspArgleuGluAspSerAsnLysleuYrAspSerPhe 400
QY 1201 TCTCTGATTCATCAATCAGTTTCAATATTCAGCAGACATGGCAGCATCTTACATGCTTC 1260
DB 401 SerleuAspIleLysSerValSerYrSerAlAserMetAlAlAserleuHshGlyPhe 420
QY 1261 TCTCATCAGAAAGATGCTTGGAAATCTACAACTTCCACTACTCTGATGGAGCAAC 1320
DB 421 SerSerSerGluAspAlaleuGluYrYrLysProSerThrThrLeuAspGlyThrasn 440
QY 1321 AAGCTATGCCAGTTTCTCCCTGTTTCAAGAACTATGGAAGAGACTCCGGAACAGTCTT 1380
DB 441 LysleuGlyGlnPheSerProValGlnGluLeuSerGluGlnhmProGluThrSerPro 460
QY 1381 GATPAGAGAGAGCAGCATCCCAAGAAAGCTGACAGCCGCGAGCTTCCAGACGACAG 1440
DB 461 AspLysleuGluIlnAlaserIleProLysLysleuGlnThrAlaArgProSerAspSerGln 480
QY 1441 AGCAACGATTCATTCGCTGCTGCAACAGCAGAGTGGCAGCGCCAGAGTCCCTTTA 1500
DB 481 SerLysArgleuHshSerValArgThrSerSerSerGlyThrAlaglnAspSerleuLeu 500
QY 1501 TCTCCACTGATCAGAAAGTGGAGCGTGGAGAGCAATTAACAACAGCTTCTTTTCCG 1560
DB 501 SerProleuHshArgSerGlySerValGlnAspAsnYrshThrSerPheleuPheGly 520
QY 1561 CTTTTCCACGACGACAGCAGCTTCAAGAGTCTGCTGGCTGGAGCTTAAAGGCTGGCAC 1620
DB 521 LeuSerThrSerGlnGlnHshleuThrLysSerAlaglyleuGlyleuYrAspGlyTrpHis 540
QY 1621 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGC 1680
DB 541 SerAspIleleuAlarProGlnhmSerhmProSerleuThrSerSerTrpYrPheAla 560
QY 1681 ACAAGTCTTCAACTTCTACTGCTGACCATCTACGAGGAGGAGGAGGAGTACTCT 1740
DB 561 ThrGlnSerSerHisPheYrSerAlAserAlalyGlySerAlAserYrSer 580
QY 1741 GCTTACAGCTGACAGCAGCTGCCCACTTGGAGAGACCAAGTCTTATCTGTGCGAGCGG 1800
DB 581 AlalyrSerCysSerGlnleuProhmCysGlyAspGlnValYrSerValArgArg 600
QY 1801 CAGAAGCAGTACAGAGCTTACTCGCGGCGAGCTGGCAAGAGAGCCCTTTGAA 1860
DB 601 GlnYrSerSerAspArgAlAspSerArgArgSerTrpshGluGluSerProPheGlu 620
QY 1861 AAGCAGTTTAAACGCAAGAGCTGCCAAATGGAATTTGAGAGACATCATGTCAGAGAAC 1920
DB 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGluAsn 640
QY 1921 AGGTCAACGGGAGAGACTGGGAAAGTGGGAGTCAAGTCTTATGCGGACATGAGAA 1980
DB 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
QY 1981 ATCATTTGAGTCTCC 1995
DB 661 IleIleGluValSer 665

RESULT 10
ABB97291
ID ABB97291 standard; protein; 665 AA.

XX AC ABB97291;
 XX XX
 DT 27-JUN-2002 (first entry)
 XX DE Novel human protein SEQ ID NO: 559.
 XX
 KM Human, anti-infective; vulnery; anti-inflammatory; immunomodulator;
 KM anti-infective; cerebroprotective; cytoskeletal; rheumatic; gene therapy;
 KM neuroprotective; anti-Parkinsonian; protein therapy; EST;
 KM expressed sequence tag.
 XX
 OS Homo sapiens.
 PN MO200222660-A2.
 XX
 PD 21-MAR-2002.
 XX
 PF 10-SEP-2001; 2001MO-US026015.
 XX
 PR 11-SEP-2000; 2000US-00659671.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Aundl Y, Zhang J, Zhao QA, Ren F;
 PI Xue HJ, Yang Y, Wehman T, Dmanac RT;
 DR WPI; 2002-292408/33.
 DR N-PSDB; ABN32477.
 XX
 PT An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.
 XX
 PS Example 2; SEQ ID NO 559; 509pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 CC
 SQ Sequence 665 AA;
 Alignment Scores:
 Pred. No.: 3,16e-287 Length: 665
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 94.22% Indels: 0
 DB: 5 Gaps: 0
 US-10-029-345a-108_COPY_538_2532 (1-1995) x ABB97291 (1-665)
 QY 1 ATGGCCATGAGATGATGGAAGTCAATGTTACTAGAGAGTTGGTGGCTGCTGAA 60
 DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 QY 61 AGTGAACGGAAGAAAGTGTCTGCTAATTGATGAGCCGCCATTGTTGGAATCAATACCC 120
 DB 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 QY 121 CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATGAAGGAGAGTTGCAACG 180
 DB 41 HisIleuGluValAlaIleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
 QY 181 GACAAAGTTAATTAAGAGATCAATCAAGATTCAGCGAAGCAATAGTTGACATGAT 240
 DB 61 AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80

QY 241 TGCAGTCAGAGGTTAGTTACGATCAAGCTCCCAAGATGTGCTCTCTTCA 300
 DB 81 CysSerGlnIleValValValIleValIleValIleValIleValIleValIleValIleVal 100
 QY 301 GACTGTTTCTCAGTCTTCTGAGGTAACCTGAGAGAGCTTCACTCTGTTCACTG 360
 DB 101 AspCysPheLeuThrValLeuLeuGlyIleGluGluIleSerPheLeuValIleLeu 120
 QY 361 CTTCGAGGTGGTTGCTGAGTCTCTGCTGTTTCCCTGGCTCTGTGAGGAAATTC 420
 DB 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuGlyGlyIleSer 140
 QY 421 ACTCTAGCTCCATCCGCAATTTCTCCGCTGCTTACCTGTTGCAACATGGGCCAAC 480
 DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 QY 481 CGAATCTTCCCAATCTTATCTTGAGCTGCCAGCGAGATGCTCTCAACAGAGCTGATA 540
 DB 161 ArgIleLeuProAsnLeuIleuGlyCysGlnArgAspValLeuAsnIleGlyLeuMet 180
 QY 541 CAGCAGAAATGGGATGGTATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 600
 DB 181 GlnGlnAsnGlyIleGlyIleValLeuAlaAsnIleSerAsnThrCysProIleProAspPhe 200
 QY 601 ATCCCGAGTCTCATTTCTCGGCTGCTGCTGCAATGAGACGTTTGTGAGAAATTTG 660
 DB 201 IleProGlnSerIlePheLeuArgValProValAsnAspSerPheCysGlyIleLeu 220
 QY 661 CCGTGGTTGGAACAATCAGTATGATTCATTTAGAGAAAGCAAAAGCTTCAATGATGTT 720
 DB 221 ProThrLeuAspIleSerValIlePheIleGluValAlaIleAsnArgGlyCysVal 240
 QY 721 CTAGTCACTGTTTACCTGGATCTCCGCTCCGCCACATGCTATGCTTACATCATG 780
 DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
 QY 781 AAGAGATGAGACATGCTTATGATGAGAGCTTACAGTTTGTGAAAGAAAGAAAGCTACT 840
 DB 261 LeuArgMetAspPheSerLeuAspGluAlaIleArgPheValIleGluValAspProThr 280
 QY 841 ATATCTCCAAATCTCAATTTCTGGGCAACCTCTGAGTATGAGAAAGATTAAGAAC 900
 DB 281 IleSerProAsnPheAsnIleuGlyGlnIleLeuAspTyrGluIleIleValAsn 300
 QY 901 CAGACTGAGACATCAGAGGCCAAAGAGCAAACTCAAGCTCTGCACTGAGAAAGCAAT 960
 DB 301 GlnThrGlyAlaSerGlyProIleSerIleValLeuLeuHisIleGluIleProAsn 320
 QY 961 GAACCTGCTCCCTGCTCTCAAGAGGTGACAGAAAGGAGACGCCCTCACTCAACCC 1020
 DB 321 GluProValProAlaValSerGluGlyGlnIleValSerGluThrProLeuSerProPro 340
 QY 1021 TGGGCGAGCTGCTGCTCAAGAGCGAGGAGCAAAAGGCCGCTGATCCCGCAGCGTG 1080
 DB 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 QY 1081 CCCAGCGTCCAGCGCTGAGCGCTGTTAGAGAGACGCCGCTGATCAGCGCTTC 1140
 DB 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
 QY 1141 AGTGGCTGCACTGCTCCGACAGACAGGCTGGAAGACAGCAATAGCTCAAGCTTCTTC 1200
 DB 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIleValLeuValAspSerPhe 400
 QY 1201 TCTCTGATATCAATCAGTTTCAATATCAATGAGCGAGATGCGACATCTTACATGAGCTTC 1260
 DB 401 SerLeuAspIleIleSerValSerTyrSerAlaSerMetAlaIleSerLeuHisGlyPhe 420
 QY 1261 TCTCATAGAAAGATGCTTGAATATCAAAAGCTTCACTATCTGATGGAGCAAC 1320
 DB 421 SerSerSerGluAspAlaLeuGluIleTyrTyrIleProSerThrThrLeuAspGlyThrAsn 440

```
QY 1321 AAGCTATGCAAGTCTCCCTGTTGAGGAATTATCGAGCAAGACTCCGAAACAGTCC 1380
DB 441 LysLeuCyGlnPheSerProValGlnLeuSerGlnGlnThrProGlnThrSerPro 460
QY 1381 GATTAGGAGGAAGCCAGATCCCAAGAGCTGAGACCCGACGGCCCTTCAGACGCG 1440
DB 461 AspLysGlnGlnAlaSerIleProLysLeuGlnThrIleArgProSerPheSerGln 480
QY 1441 AGCAAGCATGTCATTCCGTGAGAAACAGACGAGCTGGACCGCCCAAGAGTCCCTTTA 1500
DB 481 SerLysArgLeuHisSerValArgThrSerSerSerIleThrAlaGlnArgSerLeu 500
QY 1501 TCTCACTGCAATCGAAGTGGAGCTGAGAGACATATTAACAACAGCTTCTTTGCGC 1560
DB 501 SerProLeuHisArgSerGlnSerValGlnAspArgAsnIleHisThrSerPheLeuHis 520
QY 1561 CTTTCCACCAAGCAGACAGCACTCAAGAGTCTGGCTGGCTGGCTTAAAGGCTGAGC 1620
DB 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyThrHis 540
QY 1621 TCGGATATCTTGGCCCCCAAGACTTACCCCTTCCTGACAGACAGCTGATTTTGGC 1680
DB 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerThrIlePheAla 560
QY 1681 ACAAGTCTCTCACTTCTACTCTGCTGACGACATCTTACGAGAGCAAGTCCAGTTACT 1740
DB 561 ThrGlnSerSerHisPheThrSerAlaSerAlaIleIleThrGlyLysSerAlaSerThr 580
QY 1741 GCCTAAGCTGACAGCAGCTGCTCCCACTTGGCAGAGACCAAGTCTATTTGTGCGCAGCG 1800
DB 581 AlaIleThrSerCySerGlnLeuProThrCySerGlnAspGlnValIleThrValAlaGlyArg 600
QY 1801 CAGAACCCAAAGTACAGAGCTGACTCGCGCGAGAGCTGGAGTAAGAAGAGCCCTTTGAA 1860
DB 601 GlnLysProSerAspArgAlaAspSerHisArgSerThrIleGlnLysSerProPheGln 620
QY 1861 AAGCAGTTTAAACGAGAAAGCTGCCAATATGAAATTTGAGAGAGCATCATGTCAGAGAAC 1920
DB 621 LysGlnPheLysArgArgSerCySerGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGGTACCGGAAAGCTGGGAGAAAGTGGAGCATGATGATGATTTTCCGGGCGAGTGA 1980
DB 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerPheSerIleSerMetGln 660
QY 1981 ATCATTGAGGTCTCC 1995
DB 661 IleIleGlnValSer 665
RESULT 11
ADA54744
ID ADA54744 standard; protein; 665 AA.
XX
AC ADA54744;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 2312.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane protein; cancer;
KW Inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PR 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
XX
PR 24-JAN-2002; 2002US-0350435P.
XX
```

```
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuhio Y,
XX
DR WPI; 2003-395539/38.
DR N-PSDB; ADA53105.
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 2312; 205bp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
XX inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 665 AA:
```

Alignment Scores:

Pred. No.:	1.28e-286	Length:	665
Score:	3399.00	Matches:	662
Percent Similarity:	99.70%	Conservative:	1
Best Local Similarity:	99.55%	Mismatches:	2
Query Match:	94.02%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345A-108_COPY_538_2532 (1-1995) x ADA54744 (1-665).

```
QY 1 ATGCCCATGATGATGTAAGTCAATATGTTATGTTAGAGAGTGGTGGCTTCTGCTGAA 60
DB 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnLysLeuValAlaLeuGln 20
QY 61 AGTGGACGGAAGAAAGCTGCTTAATGATGATGACCGGCAATTTGTGAAATACATC 120
DB 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnThrAsnThr 40
QY 121 CACATTTTGGAAAGCATTAATATCAACTGCTCCAACTTATGAAAGGAAGCTTGA 180
DB 41 HisIleLeuGlnAlaIleAsnIleAsnCySerLysLeuMetLysArgArgLeuGln 60
QY 181 GACAAAGTAAATTAACAGAGCTCATTCAGCATTCAGCGAAACATPAGGTTGACATGAT 240
DB 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaIleLysIleValAspIleAsp 80
QY 241 TGCAGTCAGAAAGTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTTCA 300
DB 81 CysSerGlnLysValIleValIleThrAspGlnSerSerGlnAspValAlaSerLeuSer 100
QY 301 GACTGTTTCTCATCTGACTTCTGGGTAAACTGAGAAAGACTTCAACTGTTCACCTG 360
DB 101 AspCyPheLeuThrValIleLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTGCAAGTGGGTTTGTGAGTTCCTGCTGTTTCCCTGGCCCTGTGAAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGlyPheSerArgCyPheProGlyLeuCyGlnGlyLysSer 140
QY 421 ACTCTAGTCCCTAGCTGATTTCTCAGCTTGTCTTACCTGTGCAACATGGGCCAACC 480
DB 141 ThrLeuValProThrCyHisSerGlnProCySerLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTTCCCAATCTTTATCTTGCTGCGCAGAGAGATGCTCAACAAGAGCTGAT 540
DB 161 LeuIleLeuProAsnLeuIleThrGlyCyGlnArgAspValIleAsnLysGlnLeuMet 180
QY 541 CAGCAGAAATGGATGGTATGTTTAAATGCGAGCTATACCTGTCCAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerAsnThrCyProLysProAspPhe 200
```

QY 601 ATCCCGAGTCTGATTTCTGGGTCCTGTAATGACAGCTTTGTGAGAAATTTTG 660
 Db 201 TLeProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleLeu 220
 QY 661 CCGTGGTGGCAAAATCAGTAGATTTTCATTTAGAAAAGCAAAAGCTTCAATGATGTT 720
 Db 221 ProTriPheAspIysSerValAspPheIleGluValAlaValAlaSerAsnGlyCysVal 240
 QY 721 CTAGTGACATGTTTACGTGGATCTCCCGCTCCGCCACCATCGCTATCGCTTACATCAG 780
 Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 QY 781 AAGAGATGAGCAATGCTTGAATGAGATTTGATGAGAAAAGAAAGCTACT 840
 Db 261 LysArgMetAspPheSerLeuAspGluAlaTyrArgPheValIlysgIuValArgProThr 280
 QY 841 ATATCTCAAACTTCAATTTTCTGGCCAACTCTGTGACTATGAGAAAGATTAAGAAC 900
 Db 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluValIlyValAsn 300
 QY 901 CAGATGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAAACCAAT 960
 Db 301 GlnThrGlyAlaSerGlyProIysSerTyrLeuIlyLeuLeuHisPheGluIlyProAsn 320
 QY 961 GAACCTGCTCCCTGCTCTGAGAGGTCGACAGAAAAGAGAGCCCTGACTCCACCC 1020
 Db 321 GluProValProAlaValSerGluGlyGlyGlnIlySerGlnIlyThrProLeuSerProPro 340
 QY 1021 TGTGCGGACTGCTGCTCTGAGAGGTCGACAGCAAAAGCCGCTGCATCCGCGAGCGTG 1080
 Db 341 CysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSerVal 360
 QY 1081 CCCAGCGTCGCCAGCGTCGCTGCTGTAAGAGACAGCCCGCTGTATCAGGCGCTC 1140
 Db 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
 QY 1141 AGTGGGCTGACCTGCTCCCGAGACAGGCTGGAAGACGAAATAGCTCAAGCTTCTTC 1200
 Db 381 SerGlyLeuHisPheLeuAlaAspArgLeuGluAspSerAsnIlyValArgSerPhe 400
 QY 1201 TCTCTGATATCAATCATGTTTATATTCAGCAGCATGAGCATCTTACATGCTTC 1260
 Db 401 SerLeuAspIleIlySerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 QY 1261 TCTCTCATCAGAGATGCTTGAATATCTCAAACTTCACTGATGAGGACCAAC 1320
 Db 421 SerSerSerGluAspAlaLeuGluTyrTyrIlyProSerThrThrLeuAspGlyThrAsn 440
 QY 1321 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATCGGAGAGACTCCGAAACCACTGCT 1380
 Db 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnIlyThrProGluThrSerPro 460
 QY 1381 GATTAAGAGAGAGCAGCATCCCGCAAGAGCTGACAGCCGAGGCTTACAGACAGCAG 1440
 Db 461 AspIlyGluGluAlaSerIleProIlyValPheGlnIlyThrAlaArgProSerAspSerGln 480
 QY 1441 AGCAAGCATTTGATTCGATCAGAACCCAGACAGAGTGGACCCGCAAGAGTCCCTTTTA 1500
 Db 481 SerTyrArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 QY 1501 TCTTCACTGATGAGAGTGGAGCGTGGAGAGCAATTAACACACAGCTTCTTTTCGGC 1560
 Db 501 SerProLeuHisArgSerGlySerValGluAspAsnIlyHisThrSerPheLeuPheGly 520
 QY 1561 CTTTCCACGACGACGACACTCAGCAAGTGTGCTGGCTTGAAGGCTGAC 1620
 Db 521 LeuSerThrSerGlnGlnHisIleLeuThrIlySerAlaGlyLeuIlyLeuIlyValIlyPhe 540
 QY 1621 TCGGATATCTTGAGCCCCCAGACCTTACCCCTTCCCTGACAGACAGCTGTGATTTTGGC 1680
 Db 541 SerAspIleLeuAlaProGlnIlyThrSerThrProSerLeuThrSerSerTyrIlyPheAla 560

QY 1681 ACAGACTCTTCACTTACTTCTGCTCAGCCATCTACGAGGAGGAGTCCAGTACTCT 1740
 Db 561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyIlySerAlaSerTyrSer 580
 QY 1741 GCCTACAGCTGACAGCCAGCTGCCACTTCCGAGACCAAGCTATTTCTGTGCGACGCG 1800
 Db 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
 QY 1801 CAGAACCAAGTACAGACTGACTCTGCGCGGAGAGTGCATCAAGAGAGCCCTTGA 1860
 Db 601 GlnIlyProSerAspArgAlaAspSerArgSerTyrPheIleGluIlySerProPheGlu 620
 QY 1861 AAGCAGTTTAAACGAGAGCTGCAATGAAATGAAATTTGAGAGACATGTCAGAGAAC 1920
 Db 621 LysGlnPheIlyArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
 QY 1921 AGGTCACGGAGAAAGCTGGGAGAAAGTGGCACTGACTTATGCTTTTGGGACAGATGAA 1980
 Db 641 ArgSerArgGluGluLeuGlyValGlySerGlnSerSerPheSerGlySerMetGlu 660
 QY 1981 ATCATGAGTCTCC 1995
 Db 661 IleIleGluValSer 665
 RESULT 12
 AAU79161
 ID AAU79161 standard; protein; 665 AA.
 AC AAU79161;
 DT 02-JUL-2002 (first entry)
 DE Human dual-specificity phosphatase-3 (DSP-16) mutant protein #1.
 XX Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KW cancer; graft-versus-host disease; allergy; metabolic disease;
 KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KW intercellular adhesion; DSP-16 modulator; mutant; mutagen.
 XX Homo sapiens.
 OS Synthetic.
 FT Key Location/Qualifiers
 FT Misc-difference 213 /note= "Wild-type Asp substituted by Ala"
 PN MO200226997-A2.
 XX 04-APR-2002.
 PD 25-SEP-2001; 2001WO-US030124.
 PF 26-SEP-2000; 2000US-0235487P.
 PR (CEPT-) CEPTYR INC.
 PA Luche RM, Wei B;
 XX MPI; 2002-315802/35.
 DR New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 XX Claim 46; Page; 87pp; English.
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16

CC activity, for modulation of a proliferative response in a cell, survival
 CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present amino acid
 CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 CC mutant protein #1. Note: This sequence is not shown in the specification
 CC in figure 2 of the specification
 CC XX

Sequence 665 AA:

Alignment Scores:

Pred. No.:	1,57e-286	Length:	665
Score:	3398.00	Matches:	662
Percent Similarity:	99.70%	Conservative:	1
Best Local Similarity:	99.55%	Mismatches:	2
Query Match:	94.00%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345a-108_COPY_538_2532 (1-1995) x AAU79161 (1-665)

```

QY 1 ATGGCCCATGAGATGATGGAATCTCAATTGTTCTGAGAGGTGGCTGCTGAGAA 60
DB 1 MetAlaHisGluMetCileGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAAAAGTCTGCTAATTGATGCGCGGCATTTGGGAATCAATACATC 120
DB 21 SerGlyThrGluValLeuLeuIleAspSerArgProheValGluTyrsnThrSer 40
QY 121 CACATTTTGAAGCCATTAATATCAACTGCTCCAGCTTATGAAGGAAAGTTGCAACG 180
DB 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIysLeuMetIysArgIleuGln 60
QY 181 GACAAAGTTTAATTAACAGACTCATCCAGCTTTCAGCGAAACATTAAGTTGACATTGAT 240
DB 61 AspIysValLeuIleThrGluLeuIleGlnHisSerIleAlaIysValAspIleAsp 80
QY 241 TGCAGTGAAGAGGTGTAGTTACGATCAAGCTCCCAAGATGTTGCCCTCTCTCTCA 300
DB 81 CysSerIleuValValValIysAspGlnSerSerGlnAspValAlaSerIleuSer 100
QY 301 GACTGTTTTCACCTGACTTCTGGGTAATGAGAGAGAGTTCAACTGTTCACTG 360
DB 101 AspCysheLeuThrValLeuLeuGlyIysLeuGluIysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGTGGGTTTGTGAGTTCTCGTTGTTTCCCTGCTCTGAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuCysGluGlyIysSer 140
QY 421 ACTCTAGTCCCTACCTGCACTTCTGAGCTTGTCTTACCTGTTGGCAATTTGGCCCAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CSAATTTCTCCCAATCTTATTTGGCTGCGCAAGGAGATGCTCAACAAGAGAGTGA 540
DB 161 ArgIleLeuProAsnLeuIysLeuGlyCysGlnArgAspValLeuAsnIysGluMet 180
QY 541 CAGCAAGATGGAGTTGATATGTTAAATGCCAGCTATACCTGTCAAGAGCTTGA 600
DB 181 GlnGlnAsnGlyIleGlyIysValLeuAsnAlaSerAsnThrCysProIysProAspPhe 200
QY 601 ATCCCGAGTCTCAATTTCTCGCGTGGCTGTGATGACAGCTTTTGTGAGAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAlaSerPheCysGluIysIleLeu 220
QY 661 CCGTGGTTTGAACAATAGTAGATTTTCATTGAGAAAGCAAAAGCTCCCAATGAGATGTT 720
DB 221 ProThrLeuAspIysSerValAspPheIleGlyIysValAlaIysAlaSerAsnIysCysVal 240

```

```

QY 721 CTAGTCACTGTTTATGCTGGATCTCCGCTCCGACCATGCTATGCTCATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlMet 260
QY 781 AAGAGATGGAATGCTCTTTAGATGACCTTACAGATTTTGTGAAGAAAAAGCACTACT 840
DB 261 IysArgMetAspMetSerLeuAspGluAlaTyrsArgPheValIysGluIysArgProThr 280
QY 841 ATATCTCAAACTTCAATTTTCTGGCCCACTCTGAGCATATGAGAAAGATTAAGAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuAspIysGluIysIleIysAsn 300
QY 901 CAGATGAGATGAGAGGCGCAAGAGCAAACTCAAGCTGTGACCTGAGAGAGCCAAAT 960
DB 301 GlnThrGlyAlaSerGlyProIysSerIysLeuIysLeuHisIleuGluIysProAsn 320
QY 961 GAACTGTCCCTGCTGTCTCAGAGGGTGAACAGAAAGGAGAGCGCCCTCATGTCACCC 1020
DB 321 GluProValProAlaValSerIleGluIysGlnIysSerGluIysProIysSerPro 340
QY 1021 TGTGCGCACTGTGTAACCTCAGAGCGAGAGCAAGCAAAAGCCGTCATCCGCGACGCTG 1080
DB 341 CysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSerVal 360
QY 1081 CCCAGCTGCGCCAGCGTGAAGCGGTGCTGTTAGAGAGACAGCGCGCTGTACAGCGCTC 1140
DB 361 ProSerValProSerValGlnProSerLeuLeuGluIysSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGGTGAACCTGTCGCGCAGACAGGCTGTGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
DB 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluIysSerAsnIysLeuIysArgSerPhe 400
QY 1201 TCTGTGATATCAATTCAGATTTCATATTCAAGCCAGCATGGAGAGATCTTACATGGCTTC 1260
DB 401 SerLeuAspIleIysSerValSerTyserIleSerIleAlaAlaSerLeuHisGlyPhe 420
QY 1261 TCCATATCAAGAGGCTTTTGAATATCATCAAACTTCACTACTGTGATGGAGCAAC 1320
DB 421 SerSerSerGlnAspAlaLeuGluIysIysProIysProSerThrThrLeuAspGlyThrAsn 440
QY 1321 AAGCTATGCCAGTTTCTCCCTGTTCAAGAACTATCGAGAGAGACTCCGGAACAGTCTCT 1380
DB 441 IysLeuCysGlnPheSerProValGlnGlnLeuSerGlnIleThrProGluThrSerPro 460
QY 1381 GATAAGAGAGAGAGAGATCCCAAGAACTGAGACCGCGAGCTTCAAGACAGCCAG 1440
DB 461 AspIysGluGluAlaSerIleProIysIysLeuGlnThrAlaArgProSerAspSerGln 480
QY 1441 AGCAAGCATTTGATTTGGTCAAGACAGAGAGAGAGAGAGCGCCAGAGGTCCTTTTA 1500
DB 481 SerIysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
QY 1501 TCTCCATGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB 501 SerProIleuHisArgSerGlySerValGlnAspAsnIysHisThrSerPheLeuPheGly 520
QY 1561 CTTTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB 521 LeuSerThrSerGlnGlnHisIleuThrIysSerAlaGlyLeuGlyLeuIysGlyIysHis 540
QY 1621 TCGATATCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerIysIysPheAla 560
QY 1681 ACAAGATCTCACTTCACTTGTGCTCAGCAAGATGAGAGAGAGAGAGAGAGAGAGAG 1740
DB 561 ThrGlnSerSerHisPheTyserAlaSerAlaIleIysGlyIysSerAlaSerIysSer 580
QY 1741 GCTACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
DB 581 AlaTyrsSerCysSerGlnLeuProThrCysGlyAspGlnValIysSerValAlaArgAlaGag 600
QY 1801 CAGAAAGCAAGTGAACAGAGTGAAGTCCGCGAGAGCTGAGATGAAGAGAGAGAGAGAG 1860

```

|||||
Db 601 Glnlyseroserapargalaaapserargsertrphtsglunuserpompeglu 620
Qy 1861 AAGCAGTTTAAAGCAGACAGCTGCAATGGAATTTGGAGAGCATCATGTCAAGAAC 1920
Db 621 LysGlnpheylsargargsercysglnmecGlnpheGlylusexilemwetsergluasn 640
Qy 1921 AGGTCAAGGAGAGAGCTGGGGAAGTGGGAGAGCTGAGCTTTTCGGGACAGATGGAA 1980
Db 641 ArgserargglunleuglylvalglyserGlnserPheSerGlyserMetglu 660
Qy 1981 ATCATTGAGCTCTCC 1995
Db 661 IleIleGlnvalser 665
RESULT 13
AAU79162
ID AAU79162 standard; protein; 665 AA.
XX AAU79162;
AC
XX 02-JUL-2002 (first entry)
XX
XX Human dual-specificity phosphatase-3 (DSP-16) mutant protein #2.
DE
XX Human: dual-specificity phosphatase-3; DSP-16; autoimmune disease;
XX mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
XX cancer; graft-versus-host disease; allergy; metabolic disease;
XX abnormal cell growth; abnormal cell proliferation; contact inhibition;
XX cell cycle abnormality; anchorage independent cell growth; apoptosis;
XX intercellular adhesion; DSP-16 modulator; mutant; mutain.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 244 /note="Wild-type Cys substituted by Ser"
PN MO200226997-A2.
XX 04-APR-2002.
XX 25-SEP-2001; 2001MO-US030124.
XX 26-SEP-2000; 2000US-0235487P.
XX (CEPT-) CEPTYR INC.
XX Luche RM, Wei B;
XX WPI; 2002-315802/35.
XX
XX The present invention relates to a new polypeptide, DSP-16, having a 665
XX amino acid sequence, given in the specification, or a variant having at
XX least 50 % identical residues, which retains the ability to
XX dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
XX invention can be used for identifying agents which modulate DSP-16
XX activity, for modulation of a proliferative response in a cell, survival
XX of a cell, or differentiation of a cell. The cell displays contact
XX inhibition of cell growth or anchorage independent growth and may display
XX altered intercellular adhesion. The agent may modulate apoptosis, or the
XX cell cycle. The identified modulators can be used to treat Duchenne
XX muscular dystrophy, cancer, graft-versus-host disease, autoimmune
XX diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
XX cell proliferation, and cell cycle abnormalities. The present amino acid
XX sequence represents the human dual-specificity phosphatase-3 (DSP-16)

CC mutant protein #2. Note: This sequence is not shown in the specification
CC but is derived from the wild-type human DSP-16 (AAU79162) protein given
CC in figure 2 of the specification
XX
XX SQ Sequence 665 AA;
Alignment Scores:
Pred. No.: 2,34e-286 Length: 665
Score: 3396.00 Matches: 662
Percent Similarity: 99.70% Conservative: 1
Best Local Similarity: 99.55% Mismatches: 2
Query Match: 93.94% Indels: 0
DB: Gaps: 5
US-10-029-345A-108_COPY_538_2532 (1-1995) x AAU79162 (1-665)
Qy 1 ATGGCCCATGAGATGATGGAATGGAATGGAATTTTACTGAGAGGTGTGCTGTGGAA 60
Db 1 MetAlaHieGluwetiIleglyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGlu 20
Qy 61 AGTGAACGGGAAAAAAGTCTGCTCAATTGATAGCCGGCCATTGCTGTGAATTCATATAC 120
Db 21 SerGlyThrGlnValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Qy 121 CACATTTTGAAGCCATTATATATCACTGCTCCAACTTATGAAGGAAAGTTGCAACAG 180
Db 41 HieIleLeuGlnAlaIleLeuIleLeuIleLeuIleLeuIleLeuIleLeuIleLeuIleLeu 60
Qy 181 GACAAAGTGTATATTAACAAGCTCATTCAGCATTCAGCGAACAATAGTTGACATTGAT 240
Db 61 AspLyValLeuLeuIleThrGlnLeuIleGlnHieSerAlaIleHieIleValAspIleAsp 80
Qy 241 TGCAGTCAGAGGTTGTATGATTACGATCAAGCTCCCAAGATGCTCTCTCTTCA 300
Db 81 CysSerGlnLysValValValValValValValValValValValValValValValValVal 100
Qy 301 GACTGTTTCTCACTGATCTTCTGGGTAATCGAAGAGCTTCAACTGTTCACCTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHieLeu 120
Qy 361 CTTCAGAGTGGGTTGCTGAGTTCTGCTGTTTCTGCGCTGTGAGGAAATTC 420
Db 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyLysSer 140
Qy 421 ACTTAGTCCCTACCTGCACTTCTCAGCTTGTGTTACTGTTGCAACATTTGGCCAAAC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Qy 481 CGAATCTTCCCAATCTTATCTTGCTGCTGCGAGAGATGCTCTCAACAAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuIleuGlyCysGlnArgAspValLeuAsnLysGlnLeuMet 180
Qy 541 CAGCAGAAATGGAGATTGTTATGTGTTAAATGCCAGATATACCTGCAAGCTGACTTT 600
Db 181 GlnGlnAsnGlyIleGlyLysValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
Qy 601 ATCCCGAGTCTATTTCTGCGCTGCTGCTGCAATGACAGACTTTTGTGAGAAATTTTG 660
Db 201 IleProGlnSerIlePheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
Qy 661 CGGTGTTGAGCAAAATCAGATGATTGATTGAGAAAGCAAAAGCTTCAATGATGTGTT 720
Db 221 ProThrLeuAspLysSerValAspPheIleGlnLysAlaIleAsnArgGlyCysVal 240
Qy 721 CTAGTCACTGTTTACCTGGATCTTCGCTCGCCAGCAGATGCTTATGCTTACATCATG 780
Db 241 LeuValHieSerLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIle 260
Qy 781 AAGAGATGAGACATGCTTTAATGATGAGACTTACAGATTGTGAAGAAAAAGCTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaIleLysPheValIleGlnLysAspProThr 280
Qy 841 ATATCTCCAAACTTCAATTTTCTGGGCAATCTCTGACTATGAGAAAGATTAAGAAC 900

Query Match: 93.72% Indels: 0
DB: 5 Gaps: 0
US-10-029-345a-108_copy_538_2532 (1-1995) x ABR52352 (1-665)

```
QY 1 ATGGCCATAGATGATTGGAACTCAAAATTGTTACTAGAGAGTTGGTGGCTGTCTGGAA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
QY 61 AGTGAACGGAAAAAGTGTCTTAATTGATAGCCGCGCATTTGGAAATCAATATACATCC 120
DB 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyraThrSer 40
QY 121 CAAATTTGGAACCATTAATATCACTGCTCAAGTTTGAAGCGAAGTTGGCAACG 180
DB 41 HisIleLeuGlnAlaIleAsnIleAsnCySerIleuMetIleAspArgLeuGlnGln 60
QY 181 GACAAAGTGAATTAATACAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATGAT 240
DB 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY 241 TGCAGTCAGAAAGTTGATTTACGATCAAAAGCTCCCAAGATGTGCTCTCTTCA 300
DB 81 CySerGlnLysValValValTyraAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTTCTCACTGATCTTCTGGGTAACTGGAGAAAGCTTCAACTGTCTTCACTG 360
DB 101 AspCyAspPheLeuThrValIleuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGTGGGTTTGCTGAGTTCTCTCGTTGTTTCCCTGGGCTCTGTGAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGluPheSerArgCyPheProGlyLeuCyGlnGlyLysSer 140
QY 421 ACTTGAATCCCTACCTGAGATTTCTCAGCCTTGTCTTACCTGTGCGCAACATTTGGCCAAC 480
DB 141 ThrLeuValProThrCyHisSerGlnProCyLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTTATCTTGGGTGCGACGAGATGTCCTTCAACAAAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuTyrlLeuGlyCyGlnArgAspValLeuAsnLysGluLeuMet 180
QY 541 CACGAGAATGGATGGTATGTTTAAATGCGACGATATACCTGTCCAAAGCCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyTyrlValLeuAsnAlaSerAsnThrCyAspLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGCTGCGCTGTGAATGACAGCTTTGTGAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCyGlnLysIleLeu 220
QY 661 CCGTGGTTGGACAAATCAGTAAATTTTATTGAGAAAGCAAAAGCCTTCAATGATGTT 720
DB 221 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAsnGlyCyVal 240
QY 721 CTAGTCACTGTTAGTGGATCTCCGCTCGCGCACATCGCTATCGCTCATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlIleMet 260
QY 781 AAGAGATGAGACATGTTTATGATGAAGCTTACAGATTTTGAAGAAAAAGACCTACT 840
DB 261 LysArgMetAspMetSerLeuAspGlnAlaTyrlArgPheValLysGlnLysAspProThr 280
QY 841 ATATCTCAAACTTCAATTTTCTGGGCGCACTCTGACATATGAGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAlaTyrlGluLysLysIleLysAsn 300
QY 901 CAGACTGAGACATCAGGGCGAAAGACAACTCAAGTGTCTGACCTCGAGAAAGCCAAAT 960
DB 301 GlnThrGlyLysSerGlyProLysSerLysLeuLysLeuProLeuGlnLysProAsn 320
QY 961 GAACTGTCTCTGTCTCAGAGGCTGACAGAAAAAGCAGACGCGCTCACTCCACCC 1020
DB 321 GluProValProAlaValSerGlnGlyGlnLysSerGlnThrProLeuSerProPro 340
```

```
QY 1021 TGGCCGACTCTGCTACTCAGAGGACGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1080
DB 341 CyAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal 360
QY 1081 CCCAGCGTCCCGCAGCGTGCAGCCGCTGTTAGAGAACAGCCCGTGTACAGCGCTC 1140
DB 361 ProSerValProSerAlaGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
QY 1141 AGTGGCTGCACTGTCCCGACAGACAGCTGGAAAGACAAATTAAGCTTCAAGCTTCTTC 1200
DB 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysAspSerPhe 400
QY 1201 TCTCTGATATCAAACTCAGTTTATATATCAGCCAGATGGCAGACCTTACATGGCTTC 1260
DB 401 SerLeuAspIleLysSerValSerTyrlSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
QY 1261 TCTCTCAGAAAGATCTTTGGAATCTACAAACCTTCCACTACTGTGATGGAGCCAAC 1320
DB 421 SerSerSerGlnAspAlaLeuGlnTyrlTyrlLysProSerThrThrLeuAspGlyThrAsn 440
QY 1321 AAGCTATGCCAGTTCTCCCTGTTCAGAAACTATCGAGCAGACATCCGAAACCATGCTT 1380
DB 441 LysLeuCyGlnPheSerProValGlnGlnLysSerGlnGlnThrProGlnThrSerPro 460
QY 1381 GATTAAGAGGAGCCAGCATCCCGCAAGAGCTGCAGACCGCCAGGCTTCAGACAGCCAG 1440
DB 461 AspLysGlnGlnLysSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
QY 1441 AGCAAGCGATTTGATTCGGTCAAGAACAGCAGCAGTGGCAGCCGCGAAGTCCCTTTTA 1500
DB 481 SerLysArgLeuHisSerValAlaArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
QY 1501 TCTCCACTGATGGAAGTGGAGCGTGGAGAGACAAATTAACACACAGCTTCTTTTCCGCG 1560
DB 501 SerProLeuHisArgSerGlySerValGlnAspAsnTyrlHisThrSerPheLeuPheGly 520
QY 1561 CTTTCCAGCAGCAGAGACACCTCAGAAAGTCTGCGGCTGGGCTTAAAGGCTGGCAC 1620
DB 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrlPheHis 540
QY 1621 TCGGATATCTTGGCCCCCGCAGACCTTACCCCTTCCCTGACACAGACCTGTATTTTGGC 1680
DB 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrlPheAla 560
QY 1681 ACAAGTCTCTACATTTACTCTGCTCAGCCATTAAGAGAGCAGTGCAGTTACTCT 1740
DB 561 ThrGlnSerSerHisPheTyrlSerAlaSerAlaIleTyrlGlyLysSerAlaSerTyrlSer 580
QY 1741 GCCTACAGCTGCAGCGAGCTGCCCATCTTGCGGAGCAAGTCAATTCTGTGCGCAGGCGG 1800
DB 581 AlaTyrlSerCySerGlnLeuProThrCyGlyAspGlnValTyrlSerValAlaArgArgArg 600
QY 1801 CAGAAACCAAGTACAGAGCTGACTGCGCGCGAGAGCTGGCATGAAGAGAGCCCTTTGAA 1860
DB 601 GlnLysProSerAspArgAlaAspSerAlaArgArgSerTyrlHisGlnGlnSerProPheGln 620
QY 1861 AAGCAGTTTAAACGCAAGAGCTGCCAAATGGAATTTGAGAGAGCATATGATGAGGAAC 1920
DB 621 LysGlnPheLysArgArgSerCyGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGGTCACGGGAAGAGCTGGGGAAGAGTGGGAGTCACTTACGTTTCTGGGACAGATGGA 1980
DB 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCATTTAGAGTCTCC 1995
DB 661 IleIleGlnValSer 665
```

RESULT 15
AAB20325 ID AAB20325 standard; protein; 666 AA.
XX AAB20325; AC AAB20325;

XX	29-MAY-2001	(first entry)	
DT	XX	Human protein phosphatase and kinase protein-4.	
DE	XX	Human protein phosphatase and kinase protein-4.	
KW	XX	gastrointestinal disorder; immune system disorder; neurological disorder	
KX	XX	cell proliferative disorder; cancer; diagnosis; therapy.	
OS	XX	Homo sapiens.	
FT	FT	Key	Location/Qualifiers
FT	FT	Modified-site	12
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	21
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	23
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	38
FT	FT	Modified-site	/note= "N-glycosylated"
FT	FT	Modified-site	38
FT	FT	Modified-site	/note= "N-glycosylated"
FT	FT	Modified-site	49
FT	FT	Modified-site	/note= "N-glycosylated"
FT	FT	Modified-site	72
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	82
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	85. .298
FT	FT	Modified-site	/note= "VH1-type dual specificity phosphatase signature"
FT	FT	Modified-site	91
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	190
FT	FT	Modified-site	/note= "N-glycosylated"
FT	FT	Modified-site	212
FT	FT	Modified-site	/note= "N-glycosylated"
FT	FT	Modified-site	214
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	220. .280
FT	FT	Modified-site	/note= "tyrosine specific protein phosphatase"
FT	FT	Modified-site	237. .278
FT	FT	Modified-site	/note= "Y phosphatase signature"
FT	FT	Modified-site	266
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	280
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	300
FT	FT	Modified-site	/note= "N-glycosylated"
FT	FT	Modified-site	369
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	393
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	421
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	422
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	434
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	439
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	468
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	471
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	479
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	528
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	590
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	597
FT	FT	Modified-site	/note= "O-phosphorylated"
FT	FT	Modified-site	605

FT		/note= "O-phosphorylated"
FT	Modified-site	610
FT		/note= "O-phosphorylated"
FT	Modified-site	613
FT		/note= "O-phosphorylated"
FT	Modified-site	618
FT		/note= "O-phosphorylated"
FT	Modified-site	628
FT		/note= "O-phosphorylated"
FT	Modified-site	641
FT		/note= "N-glycosylated"
FT	Modified-site	643
FT		/note= "O-phosphorylated"
XX		
PX	WO200120004-A2.	
PD		
XX	22-MAR-2001.	
PF	14-SEP-2000; 2000MO-US025515.	
PR		
XX	15-SEP-1999; 99US-0154141P.	
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
PI	Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y,	
XX	Lu DM;	
DR	WPI; 2001-244811/25.	
DR	N-PSDB; AAF30479.	
XX		
PT	Novel human protein phosphatase and kinase proteins for diagnosis,	
PT	treatment and prevention of gastrointestinal, immune system, neurological	
PT	and cell proliferative disorders.	
PS	Claim 1; Page 87-88; 103pp; English.	
CC		
CC	The present sequence is that of novel human protein phosphatase and	
CC	kinase protein PPHK-4, as predicted from incyte clone ID No. 1234795CBI	
CC	(see AA30479). Tissues that express PPHK-4 (as a fraction of total	
CC	tissues expressing PPHK-4) include gastrointestinal (0.385),	
CC	cardiovascular (0.154), nervous (0.154) and reproductive (0.154).	
CC	Diseases or conditions associated with tissues expressing PPHK-4 (as a	
CC	fraction of total tissues expressing PPHK-4) include cancer (0.692),	
CC	inflammation or trauma (0.308) and cell proliferation (0.231). The	
CC	encoded protein shows homology to mouse neuronal tyrosine threonine	
CC	phosphatase 1. The invention provides human PPHK-1 to -11 polypeptides	
CC	(see AAB20322-32) and polynucleotides (see AA30476-86). It also provides	
CC	expression vectors, host cells, antibodies, agonists and antagonists, as	
CC	well as methods for diagnosing, treating or preventing disorders	
CC	associated with expression of PPHK, including gastrointestinal	
CC	disorders, immune system disorders, neurological disorders and cell	
CC	proliferative disorders, including cancer	
XX		
SQ	Sequence 666 AA:	
	Alignment Scores:	
	Pred. No.: 6,4e-285	
	Score: 3379.50	
	Percent Similarity: 99.40%	
	Best Local Similarity: 99.25%	
	Query Match: 93.49%	
	DB: 4	
	Gaps: 1	
US-10-029-345A-108_COPY_538_2532 (1-1995) x AAB20325 (1-666)		
OY	1 ATGGCCCATGAGATGATTGGAACTCAATTGTACTGAAGGTTGCGTCTGCTGGA 60	
Db		
	1 MetAlahstsgUmetIleGlyThrGlnIlleValThrGluArgLeuValAlaIleuGlu 20	
OY	61 AGTGGAACGAAAAAGTCGTCTTAATGARGCGGCCCACTTGTCGGAATCAATCACCC 120	
Db		
	21 SerGlyThrGluIysValIleuIleuIleAspSerArgProIleValGluIysTranThrSer 40	

OY	121	AACATTTTGAAGGCATTAATATCACTGCTCAAGCTTATGAAAGCGAAGTTGCAACAG	180
Db	41	HisIleuGluAlaIleAsnIleAsnCySerIySueMetLysArgIleGln	60
OY	181	GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT	240
Db	61	AspIySValIleuIleThrGluIleuIleGlnHisSerIAlaYbHisIySValAspIleAsp	80
OY	241	TGCAGTCAAGAGGTGTATGTTACGATTAAGCTCCCAAGATGTGGCTCTCTCTTCA	300
Db	81	CysSerGlnIySValIyValIyLysAspIleSerSerGlnAspValAlaSerIeIySer	100
OY	301	GACGTGTTTTCACATGTACTCTGGGTAACTGGAGAAAGACCTTCACTGTGTCAACCTG	360
Db	101	AspCySPhleuThrValIleuIleuLysIySleuGluIySserPhleAsnSerValHisIeIy	120
OY	361	CTTCGAGGTGGGTGTGCTGAGTCTCTGTTGTTTCCCTGSCCTCTGTGAAGGAAATCC	420
Db	121	LeuIleagIyGlyPhleIleagIyPhleSerIyGySPhleProGlyIyIeucySglIyIySser	140
OY	421	ACTGTAGTCCCTACCTGCATTTCTCAGCTTGCTTACCTGTGGCAACATTTGGCGCAAC	480
Db	141	ThrIeuValProThrCySleSerGlnProCySleuProValAlaAsnIleGlyProThr	160
OY	481	CGAATTCCTCCCATCTTATCTTGCTGGCGCCAGAGATGTCCATACAAGAAGACTGATA	540
Db	161	ArgIleIeupProAsnIeulIeudIyCySglnIyArgPValIleuAsnIySglIyIeumet	180
OY	541	CAGCAGAATGGAGATGGTATGTATGTGTTAAATCCAGCTTACCTGTGCCAAAGCTGACTTT	600
Db	181	GlnGlnAsnGlyIleGlyTyValIleuAsnAlaSerAsnThrCySProIySProAspPhe	200
OY	601	ATCCCGGAGTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG	660
Db	201	IleProGlnSerHisPheIeulIyArgValProValAsnAspSerPheCySglIyIleu	220
OY	661	CCGTGCTGTGCAAAATCACTAGATTTCAATTGAGAAAGCAAAAGCTCCATAGATGTCTT	720
Db	221	ProThrIeupAspIySserValAspPheIleGluIySAlaIySAlaSerAsnGlyCySVal	240
OY	721	CTAGTGCAGCTGTTAGTGGGTCTCCGCGTCCGCGCACATCGCATGCGCTACATCAATG	780
Db	241	LeuValHisCySleuAlaGlyIleSerIySerIAlaThrIleAlaIleAlaIyIleIeuc	260
OY	781	AAGAGGATGACATGTCTTTAGATGAACCTTACAGATTTGTGAAGAAGATTAAAGAC	840
Db	261	LysArgMetAspMetSerIeupIySAlaIyAlaTyArgPheValIySglIyIyArgProThr	280
OY	841	ATATGTCCAAACTTAATTTCTGGGCGCAACTCTGTGACTTGAAGAAGATTAAAGAC	900
Db	281	IleSerProAsnPhleAsnPhleuGlyGlnIeulIeupIyTyGluIySlyIleIySAsn	300
OY	901	CAGACTGGAGATCAAGGCGCAAGGCAACTCAAGCTGCGACTGGAGAAAGCGAAAT	960
Db	301	GlnThrGlyAlaSerGlyProIySserIySleuIySleuHisIeGlnIySProAsn	320
OY	961	GAACCTGTCCCTGCTCTCAGAGGSGTGAAGAAAGCAGAGCGCCCTCAGTCCACC	1020
Db	321	GluProValProAlaValaIleSerGluIyGlyIyGlnIySserGluThrProIeuserProPro	340
OY	1021	TGTGCGCACTGTGTAACCTCAGAGGACAGACAGCAAAAGCCGTGTCATCCCGACGCGT	1080
Db	341	CysAlaAspSerIAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSerVal	360
OY	1081	CCGACGCGCCAGCGTGCAGCGCTGTAGTAGGACAGCCGCGTGTGACAGCGCTC	1140
Db	361	ProSerValProSerValGlnProSerIeueuGlnAspSerProIeValGlnIleu	380
OY	1141	AGTGGGCTGCACCTTCCGACAGACAGGCTGGAGAGACACAATTAAGCTCAAGCTTCTTC	1200
Db	381	SerGlyIeueHisIeIeSerIAlaAspArgIeueGlnAspSerAsnIySleuIyArgIySerPhe	400
OY	1201	TCTCTGATATCAAAATCAGTTTCATATTCAGCCAGCATGGACATCTCTTACATGGCTTC	1260

Db		401	SeriesunappliedeyservalsectySerslasernechallaserleuHlgllyPhe	420
Oy		1261	TCCTCATGAGAAAGTCGTTTGGAAATACTACAAACCTTCCACTACTCTGATGGGACCAAC	1320
Db		421	SerserserglunapralaleuGlunlyrTylylaspProserThrlleuAaprglyThrashn	440
Oy		1321	AAAGTAAAGCCAGTTTCCCCCTGTTCCAGGAACCTATCCGAGGAGAGACTCCCGAAAC--AGT	1377
Db		441	LysleuCysglnPheSerProvalGlnleuSerglunlntmrProglunthSerser	460
Oy		1378	CCTATATAAGGAGGAGAGACGATCCCAAGAGCTGAGACCGCAGGCTTCAGACAGC	1437
Db		461	-LeuIIlaIrglrglysrProalaserProalgserCyArgProProgllyleuGlnthral	480
Oy		1438	CAGAGCAAGCGATTTGCATTCCGTCAGAACACGACGACGAGTGGCACCCGCCAGAGGCTCCCTT	1497
Db		480	aArglaSersapryslleargserglunProalaleValaIaIaProProarglylProph	500
Oy		1498	TTATCTTCACCTGCATCGAAAGTGGAGCGGTGGAGAGCAATTACACACACAGCTTCCTTTTC	1557
Db		500	etyrleuHlascyslleGluValGlyalrPatrgThrlleThrProalaserPheSe	520
Oy		1558	GGCGCTTTCCACCGGCCAGAGACCTACGAGAGCTGCTGGCCCTGGGC--CTTAAAGGAGCTG	1616
Db		520	rAlaPheProProalaserSerThrsrAargserleuLeuAlaTrpAlaleuuybglyTr	540
Oy		1617	GCACCTCGGAATATCTGGSCCCCCAGACCTCTACCCCTCCCTGACACGAGCTGGTATTTT	1676
Db		540	pHisSerHaprlleleuAlaProglntmrSerrThrProSerleuHtserserrTrpyrph	560
Oy		1677	TGCCACAGAGTCCTCACACTTTTACTCTGCTCCAGCCATTTACGAGGACGAGTGCAGTTA	1736
Db		560	eAlaThrglunSerSerHlshPheTyrserrAlaaserAlaIleTyrclylGlyserAlaserTy	580
Oy		1737	CTCTGCGCTTAAGCTGCGACGAGCTGCCACTTGGGGAGACCAAGTCATTTCTGTGGCGCAG	1796
Db		580	rSerAlaTyserTysserserglneupProtmrCybglyAspGlnAlaTyrserrAlaAargr	600
Oy		1797	GGCGCAGAAAGCAAGTACAGAGCTGACTGGCGCGAGCTGGCATGAAGAAGACCCCTT	1856
Db		600	gArgGlnlysrProserHaprglaaapSerHrAargserTrpHlselunlunserProph	620
Oy		1857	TGAAGACCAAGTTTAAACGACGAAAGCTGCCAATGGAAATTTGGAGAGAGCATCATGTGAGA	1916
Db		620	eglunlyGlnPheHysarglrgserCysglmerGlnPheGlyglunserllemetserGlg	640
Oy		1917	GAACAGCTCACGGGAAAGAGCTGGGGGAAAGTGGGAGTCACTTACGTTTTCGGGACACAT	1976
Db		640	uAsnArgSerHrArglunlunleuGlylyyValGlySerglunserSerPheSerserglySerme	660
Oy		1977	GGAAATCATTGAGGCTCC	1995
Db		660	rglnlrlleGlnValser	666

Search completed: June 21, 2004, 12:34:14
Job time : 154.589 secs

the page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:26 ; Search time 15.7715 Seconds
(without alignments)
11051.549 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 1627
Sequence: 1 atgagccatgagatgattg9.....agaagataagaaccagacac 906

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=x1p
-Q=/cgn21/USPTO.spool_p/US10029345/runat_21062004_122816_4205/app_query.fasta_1.3278
-DB=pir_78 -QMT=fastcan -SUFFIX=xpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPM=prc -NOR=exc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US10029345.QCGN.1.1_78@runat_21062004_122816_4205 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	450	27.7	619	2	T15969
2	425	26.1	367	2	T15969
3	423	26.0	314	2	S24411
4	418	25.7	314	2	A56115
5	415	25.5	367	1	A57126
6	415	25.5	367	1	S29090
7	409	25.1	314	2	S52265
8	377.5	23.2	314	2	B57126
9	351	21.6	393	2	A56947
10	325	20.0	384	1	T13890
11	289	17.8	303	2	T16405
12	223.5	13.7	365	2	T132494
13	222.5	13.7	330	2	T139698
14	212.5	13.1	186	2	T16056
			226	2	T21380

15	208	12.8	272	2	T18915	hypothetical prote
16	207.5	12.8	220	2	JC7885	low-molecular-mass
17	205	12.6	223	2	T149365	protein tyrosine p
18	202	12.4	354	1	S31304	protein-tyrosine-p
19	201.5	12.4	142	2	T03074	dual specificity p
20	195.5	12.0	185	1	A47196	dual specificity p
21	195	12.0	283	2	G84458	probable protein p
22	193	11.9	205	2	T49364	protein tyrosine p
23	185	11.4	807	1	S44538	probable protein-p
24	183.5	11.3	276	2	T48806	protein-tyrosine-p
25	181.5	11.2	278	2	T39517	dual-specificity M
26	179.5	11.0	866	2	F88481	protein C16a3.1 (1
27	173.5	10.7	204	2	T17802	hypothetical prote
28	171	10.5	489	1	S58725	phosphatase-like p
29	162	10.0	771	2	T47666	hypothetical prote
30	160.5	9.9	580	2	T18439	hypothetical prote
31	160.5	9.9	600	2	T18446	hypothetical prote
32	157	9.6	209	1	S48459	probable dual spec
33	152	9.3	292	2	S41012	hypothetical prote
34	151.5	9.3	597	1	S43743	probable dual spec
35	150.5	9.3	169	2	T30684	probable dual spec
36	145.5	8.9	171	1	T36845	dual specificity p
37	145.5	8.9	171	2	T28522	probable dual spec
38	145.5	8.9	171	2	B72161	dual specificity p
39	144.5	8.9	171	1	Q0VZHI	dual specificity p
40	138.5	8.5	171	1	A42514	hypothetical prote
41	134	8.2	150	2	T21489	hypothetical prote
42	134	8.2	272	2	T19418	hypothetical prote
43	127.5	7.8	171	1	B47452	dual specificity p
44	126	7.7	1132	2	T49403	related to protein
45	121	7.4	928	2	S50578	hypothetical prote

ALIGNMENTS

RESULT 1

T15969
hypothetical protein F08B1.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C/Accession: T15969

R/Chisoe, S.
submitted to the EMBL Data Library, July 1995

A/Description: The sequence of C. elegans cosmid F08B1.

A/Reference number: Z16439

A/Accession: T15969

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-619 <CHI>

A/Cross-references: EMBL:U23178; NID:G726421; PID:G726422; PIDN:AC46719.1; CESP:F08B1.1

A/Experimental source: strain Bristol N2.

A/Gene: CESP:F08B1.1

A/Introns: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3

Alignment Scores:

Pred. No.: 5.25e-32 Length: 619
Score: 450.00 Matches: 90
Percent Similarity: 69.6% Conservative: 34
Best Local Similarity: 50.56% Mismatches: 48
Query Match: 27.66% Indels: 6
DB: 2 Gaps: 4

US-10-029-345A-108_COPY_538_1443 (1-906) x T15969 (1-619)

QY 367 GGTGGGTTTCTAGTCTCTCGTTTTCCTGCGCTCTGTGAAGAAATCC----- 420
DB 99 GYGLYPheylsGlnPhealagInGlyrProGlnLeuysGlnserGlnGlyMet 118
QY 421 ACTCTAGCCCTACCCGACATTCTCAGCCCTTGTGTTA-----CCTGTGCCAACATTGGG 474
DB 119 ThrArgLeuProGlnserLeuSerGlnProCyGLeuSerGlnProThrGlyAsp---Gly 137


```

QY 385 TCTGTTGTTCCCTGCGCTCTGT---GAAGGAAAATCACTTAGTCCCT----- 432
Db 136 GlnGlyCysCysProAspLeuCySerGlnAlaProAlaProAlaLeuProIleThrGly 155
QY 433 -----ACCTGCAATTTCTGAGCCTTGCTTACCTGTGGCAACATT---GGGCCAACCAG 483
Db 156 AspLysThrSerArgSerAspSerArgAlaProValIleThrSerGlnGlyProValGlu 175
QY 484 ATTCTTCCCAATCTTTATCTTGTGCTGCTCCAGCAAGATGCTCTCAACAAGAGCTGATACG 543
Db 176 ILeuProIleuLeuPheLeuGlySerCysSerHisSerSerAspSerGlnGlyLeuGln 195
QY 544 CAGAAATGGATTTGGTTATGTGTAAATGCGACATACCTGCTCCCAAGCCTTATTC 603
Db 136 AlaCysGlyIleThrAlaValIleuAsnValSerAlaSerCysPro----- 210
QY 604 CCCGAGTCTCATTTT-----CTCGGTGCTGCTGTGAATGACAGCTTT 645
Db 211 -----AsnHisPheGlnGlyLeuPheArgTyrIleSerIleProValGlnAspAsnGln 228
QY 646 TGTGAGAAAATTTTCCGCTGTTGACCAATACATGATTCATTCATGAGAAAACCAACC 705
Db 229 MetValGlnIleSerAlaIleThrPheGlnGlnAlaIleGlyPheIleAspTyrValIleAsn 248
QY 706 TCCAAATGATGTGTTCTAGTGACGTGTTAGCTGGATCTCCGCTCCGACCATGCGT 765
Db 249 SerGlyGlyArgValLeuValHisCysGlnAlaGlyIleSerArgSerAlaThrIleCys 268
QY 766 ATGCGCTACATCATGAAAGCATGATGCTTTAGATGAAAGCTTACATTTGTGAAA 825
Db 269 LeuAlaTyrLeuMetGlnSerArgArgValArgLeuAspGlnAlaPheAspPheValIys 288
QY 826 GAAAAAGACCTATATATCTCAAACTTCAATTTTGGGCCCACTCTGGAATGAG 885
Db 289 GlnArgArgGlyValIleSerProAsnPheSerPheMetGlyGlnLeuLeuGlnPheGln 308
QY 886 AAGAAGATT 894
Db 309 ThrGlnVal 311

RESULT 5
S29090
dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - human
M.Alternate names: protein-tyrosine-phosphatase Ctl100; protein-tyrosine-phosphatase, not
C.Species: Homo sapiens (man)
C.Date: 25-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999
C.Accession: S29090; A53052
R.Keyseq: S.M.; Emslie, E.A.
Nature 359, 644-647, 1992
A.Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine
A.Reference number: S29090; MUID:93024952; PMID:1406996
A.Accession: S29090
A.Molecule type: mRNA
A.Residues: 1-367 <KMA>
R.Keyseq: S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.
J. Biol. Chem. 269, 3596-3604, 1994
A.Title: Isolation and characterization of a human dual specificity protein-tyrosine pho
A.Reference number: A53052; MUID:94148864; PMID:8106404
A.Accession: A53052
A.Molecule type: DNA
A.Residues: 1-367 <KMA>
A.Experimental source: leukocyte
A.Note: sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804,
C.Genetics:
A.Gene: GDB:IDUSP1:PTPN10
A.Cross-references: GDB:136197; OMIM:600714
A.Map position: Sg34-Sg34
C.Superfamily: dual specificity phosphoprotein phosphatase 1; VAI-type dual specificity
C.Keywords: heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induced P
F.181-312/Domain: VAI-type dual specificity phosphoprotein phosphatase homology <VAI>
F.128/Active site: Cys (phosphocysteine intermediate) #status predicted

```

```

F.1264/Binding site: substrate phosphate (Arg) #status predicted
Alignment Scores:
Pred. No.: 6,7e-29
Score: 415.00
Percent Similarity: 49.50%
Best Local Similarity: 33.89%
Query Match: 25.51%
DB: 1
Gaps: 4
US-10-029-345a-108_copy_538_1443 (1-906) x S29090 (1-367)
QY 43 TGTGAGCTCTGCTGGAAGTGAACGAAAAGTGTCTTAATTAATGACCGGCATTT 102
Db 13 LeuArgAlaLeuLeuGlyGlyAlaGlyAlaGlnCysLeuLeuAspCysArgSerPhe 32
QY 103 GTGAAATACATATACATCCCAATTTTGAAGCATTATATATCAATGCTCCAAAGCTTATG 162
Db 33 PheAlaPheAsnAlaGlyHisIleAlaGlySerValAsnValArgPheSerThrIleVal 52
QY 163 AAGGAAAGTTGCAACAGCAACAAAGTTTAATTAACAGCTCATTCACAGTCCAGAA 222
Db 53 ArgArgArgAlaIleGlyGlyAlaMetGlyLeuGlnHisIleValProAsnAlaGlyLeuArg 72
QY 223 CATAGAGTTGACATGATTCAGACTGAAAGCTTGTGATTTAGATCAAAAGCTCCCAAGAT 282
Db 73 GlyArgLeuLeuAlaGlyAlaTyrHisAlaValIleLeuAspGlnValArgSerAlaAla 92
QY 283 GTTGCCCTCTCTCTTCAAGACTGTTTCTCACTGATCTTGGGTAACCTG---GAG 336
Db 93 LeuAspGlyAlaIleAlaArgAspGlyThrIleuAlaLeuAlaGlyAlaLeuCysArgGln 112
QY 337 AAGAGCTTCAACTGTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
Db 113 AlaArgAlaIleAlaGlnValaPhePheLeuLeuGlyGlyGlyGlnAlaPheSerAlaSerCys 132
QY 397 CCTGGCTCTGTGAAGAAA-----TCCACTCTA 426
Db 133 ProGlnLeuCysSerIleGlnSerThrProMetGlyLeuSerIleuProLeuSerThrSer 152
QY 427 GTTCCCTTACCTGATTTTCTCAGCCTTTCAGCTTTCAGCTTTCAGCTTTCAGCTTTCAGCTTTC 471
Db 153 ValProAspSerAlaGlnSerGlyCysSerCysSerThrProLeuTyrAspGlnGly 172
QY 472 GGGCCAAACCGAATTTCTTCCCAATCTTTATCTTGTGCTGCGACAGATGCTCTCAAG 531
Db 173 GlyProValGlnIleuLeuProPheLeuTyrIleGlnSerAlaTyrHisAlaSerArgLys 192
QY 532 GAGCTGATACAGCAAGATGGATGGATTAATGCTTAATGCGACTATACCTGTCCAAAG 591
Db 193 AspMetLeuAspAlaLeuGlyIleThrAlaLeuIleAsnValSerAlaAsnCysProAsn 212
QY 592 CTGACCTTATATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
Db 213 ---HisPheGlnGlyHisTyrGlnIleTyrIleSerIleProValGlnAspAsnHisIysAla 231
QY 652 AAAATTTTCCGCTGTTGCAAAATAGTAGATTATTAAGAAAGCAAAAGCTTCAAT 711
Db 232 AspIleSerSerTyrPheAsnGlnAlaIleAspPheIleAspSerIleLysAsnAlaGly 251
QY 712 GGATGTGTTTACAGCACTGTTTACGTGGATCTCCGCTCCGACCACTCCCTATGCGC 771
Db 252 GlyArgValaPheValHisCysGlnAlaGlyIleSerArgSerAlaThrIleCysLeuAla 271
QY 772 TACATCATGAAGAGATGACATGCTCTTATGATGAAAGCTTACATGCTGAAAGAAA 831
Db 272 TyrLeuMetArgThrAsnArgValIleLeuAspGlnAlaPheGlnPheValIleGlnArg 291
QY 832 AGACCTACTATCTCCAACTTCAATTTTCTGCGCACTCTGCACTATGAAAGAAAG 891
Db 292 ArgSerIleIleSerProAsnPheSerPheMetGlyGlnLeuLeuGlnPheGlnSerGln 311
QY 892 ATT 894

```



```

OY 277 CAAGATGTCCTCTCTCTTCAACCTGTTTTCACCTGACTTCTGGGTAACCTGAG 336
Db 96 ALaSerValThrGlnLeuProPheAspGlyProAlaHisLeuLeuAlaAlaLeuGln 115
OY 337 AAGAAGCTC-----AACTCTGTTCACCTGCTTGACGGGGTTCGTGAGTTC 384
Db 116 HisGlnMetArgGlyGlyProThrThrValCysPheLeuArgGlyGlyPheLeuSerPhe 135
OY 385 TCTCGTGTTCCTCCGCGCTCGT---GAAGAAATTCACCTGATCCCT-----432
Db 136 GlnThrTyrcysProAspLeucysSerGlnAlaProAlaGlnAlaLeuProProAlaGly 155
OY 433 -----ACCTGATTTCTCAGCTCTGCTTACCTGTTGCGCAACATT---GGCCCAACCGGA 483
Db 156 AlaGlnuAsnSerAsnSerAspProArgValProIleTyrcAspGlnGlyGlyProValGln 175
OY 484 ATTTTCCCAATCTTTATCTTGGCTGCCACGAGATGCTCTCAACAGAGAGCTGATACAG 543
Db 176 IleuProTyrlleuTyrlleuGlySerCysAsnHisSerSerAspLeuGlnGlyLeuGln 195
OY 544 CAGATGGGATTTGGTTATGCTTAAATGCCAGCTATACCTGTCGCAAGCTGACTTATC 603
Db 196 AlaCysGlyIleThrAlaValIleuAsnValSerAlaSerCysProAsn---HisPheGln 214
OY 604 CCGAGCTCATTTCCGCGCTGCGCTGCGTGAATGACACTTTTGTGAAAATTTTGGCCG 663
Db 215 GlyLeuPheHieTyrllysSerIleProValGlnuAsnGlnMetValGlnIleSerAla 234
OY 664 TGGTGGACAAATCGTAATTTCTTATGAGAAAGCAAAAGCTTCATGATGTTGTTCA 723
Db 235 TrpPheGlnGlnAlaIleSerPheIleAspSerValIlysAsnSerGlyGlyArgValIleu 254
OY 724 GTGCACTGTTTACCTGGGATCTCCGCTCCGCGACATCCCTATCCCTATCATCATGAAG 783
Db 255 ValHisCysGlnAlaGlyIleSerArgSerAlaThrIleCysLeuAlaTyrlleuIleGln 274
OY 784 AGATGACATGCTTTTATGATGAGCTTACAGATTTGTGAAAGAAAAGACCTACTATA 843
Db 275 SerHisArgValAlaGlnuAspGlnAlaPheAspPheValIysGlnArgGlyValIle 294
OY 844 TCTCCAAACTTCATTTTCTGGGCGCACTCTGAGTATGAGAAAGATT 894
Db 295 SerProAsnPheSerPheMetGlyGlnLeuLeuGlnLeuGlnThrGlnVal 311

RESULT 8
A56947
dual specificity phosphatase (BC 3.1.3.-) HVH2 - rat
N/Alternate names: mitogen-activated protein kinase phosphatase 2
C/Species: Rattus norvegicus (Norway rat)
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 17-Mar-1999
J/Misr-Press, A.; Rim, C.S.; Yao, H.; Roberson, M.S.; Stork, P.J.S.
A/Title: A novel mitogen-activated protein kinase phosphatase. Structure, expression, an
A/Reference number: A56947; MUID:9501550; PMID:7782322
A/Accession: A56947
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-393 <Mis>
C/Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
F/202-333/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
F/279/Active site: Cys (phosphocysteine intermediate) #status predicted
F/285/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:
Pred. No.: 1.49e-25 Length: 393
Score: 377.50 Matches: 95
Percent Similarity: 48.69% Conservative: 54
Best Local Similarity: 31.05% Mismatches: 124
Query Match: 23.20% Indels: 33
DB: 2 Gaps: 8

```

```

US-10-029-345a-108_copy_538_1443 (1-906) x A56947 (1-393)
OY 46 GTGCTCTCTGTGGAAGGTGGAACGGAAGAGGTGCTTAATGATTAACCGGCAATTGTG 105
Db 38 LeuGlyIleuLeuSerGlyGly-----LysCysLeuLeuLeuAspCysArgProPheLeu 55
OY 106 GAATACATPACATCCCAATTTTGAAGCCTTAATATCAACTGCTCAAGCTTATGAAG 165
Db 56 AlaHisSerAlaGlyTyrlleArgGlySerValAsnValArgCysAsnThrIleValArg 75
OY 166 CGAAGTTTGCAACGAGCAAGCAAGATGTTATTAAGAGCTATCCAGACTTACCGAAACAT 225
Db 76 ArgArg---AlaLysGlySerValSerLeuGlnGlnIleLeu-----ProIleGlnIleu 92
OY 226 AAGTT-----GACATGATTCAGCTCAGAAAGTTGATGATTACGAT 267
Db 93 GlnValArgProCysAlaLeuAlaSerThrArgLeuSerSerSerThrMetGlnPro 112
OY 268 CAAGCTCCCAAGATGTCCTCTCTCTGTCAGACTGTTTCTCAGCTGACTTCTGGCT 327
Db 113 AlaArgArgGlnuSerProGlyGlnHisSerValAlaGlyArgAlaAlaLeuArgArg 132
OY 328 AATCGAGAGAAGCTTCAACTGTGTCACCTGCTGACAGGTGGTTCGATGTTCTCT 387
Db 133 AsnAlaGlnuArgThr-----AspIleCysLeuLeuLysGlyGlyTyrlleArgPheSer 150
OY 388 CGTTGTTCCCTGGCTCTGTGAGAAAGAAATCACTGATCCCTACCTGACTTCTCAG 447
Db 151 SerGlnTyrlleProGlnPheCysSerIlyThrLysAlaLeu-----AlaAlaIleProPro 168
OY 448 CCTTGCTTACCTGTTGCCACATT-----471
Db 169 ProValProProSerThrAsnGlnuSerLeuAspLeuGlyCysSerSerCysGlyThrPro 188
OY 472 -----GGGCAACCCCAATTTCTCCCAATCTTATCTTGGCTGCGCAGGA 516
Db 189 LeuHisArgGlnGlyGlyProValGlnIleLeuProPheLeuTyrlleuGlySerAlaTyrl 208
OY 517 GATGCTTCGAAGAGAGCTGATACAGCAAGAAATGGATGGTATGATGTTAAATGCCAG 576
Db 209 HisAlaAlaArgArgAspMetLeuAspAlaLeuGlyIleThrAlaLeuLeuAsnValSer 228
OY 577 TATACCTGTCCAAAGCTGACTTATCCCGGAGCTCATTTCTGCGCTGCTGCAAT 636
Db 229 SerArgProAsn---HisPheGlnGlyHisTyrlleTyrlleCysIleProValGln 247
OY 637 GACAGCTTTTGAGAAATTTTGGCGGTGGACAATCAGAGATTCATTGAGAAA 696
Db 248 AspAsnHisAlaAlaAspIleSerSerThrPheMetGlnAlaIleGlnTyrlleAspAla 267
OY 697 GCAAAAGCTCCAAATGATGTTCTAGTGACGCTTTTACCTGGATCTCCGCTCGCC 756
Db 268 ValLysAspCysArgGlyArgValAlaLeuValHisCysGlnAlaGlyIleSerArgSerAla 287
OY 757 ACCATCGCTATCCCTCATCATATGAGAGATGAGCATGCTTTTACAGTAAGCTTACAGA 816
Db 288 ThrIleCysLeuAlaTyrlleuMetMetLysArgValArgLeuGlnIleuAlaPheGln 307
OY 817 TTTGGAAGAAAAGAAACCTATCATATTCGCAAACTTCAATTTTGGCGCACTCTCTG 876
Db 308 PheValLysGlnuArgArgSerIleIleSerProAsnPheSerPheMetGlyGlnLeuLeu 327
OY 877 GACTATGAGAAAGAAATT 894
Db 328 GlnPheGlnuSerGlnVal 333

RESULT 9
138890
dual specificity phosphoprotein phosphatase (BC 3.1.3.-) 5 - human
N/Alternate names: dual specificity phosphatase B23; dual-specificity phosphatase hvh-3;
C/Species: Homo sapiens (man)
C/Date: 16-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 11-Jun-1999

```

C/Accession: I38890; A55313
 R/Kwak, S.P.; Dixon, J.E.
 J. Biol. Chem. 270, 1156-1160, 1995
 A/Title: Multiple dual specificity protein tyrosine phosphatases are expressed and regulated by cytokines
 A/Reference number: A55432; MUID:95138103; PMID:7836374
 A/Accession: I38890
 A/Molecule type: mRNA
 A/Residues: 1-384 <RES>
 A/Cross-references: EMBL:U6996; NID:6642012; PIDN:AB06261.1; PID:6642013
 A/Experimental source: placenta
 R/Ishihashi, T.; Bottaro, D.P.; Micheli, P.; Kelley, C.A.; Aaronson, S.A.
 J. Biol. Chem. 269, 29897-29902, 1994
 A/Title: A novel dual specificity phosphatase induced by serum stimulation and heat shock
 A/Reference number: A55313; MUID:9505849; PMID:7961985
 A/Accession: A55313
 A/Molecule type: mRNA
 A/Residues: 1-8, 'GHV', 12-70, 'R', 72-104, 'F', 107-362, 'RCLPTQSSSAEALMQRNPARTGMEASAPQEOE
 A/Cross-references: GB:U15932; NID:9606971; PIDN:AAA64693.1; PID:9606972
 A/Experimental source: mammary epithelial cells
 C/Genetics:
 A/Gene: GDB:DUSP5
 A/Cross-references: GDB:385447
 A/Map position: 10q25-10q25
 C/Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity
 C/Keywords: phosphoprotein; phosphoric monoester hydrolase
 F/186-317/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>
 F/263/Active site: Cys substrate cysteine intermediate #status predicted
 F/269/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:
 Pred. No.: 3,44e-23 Length: 384
 Score: 351.00 Matches: 94
 Percent Similarity: 46.25% Conservative: 54
 Best Local Similarity: 29.38% Mismatches: 124
 Query Match: 21.57% Indels: 48
 Gaps: 7

US-10-029-345a-108_copy_538_1443 (1-906) x I38890 (1-384)

```

QY 52 CTGCTGAAAGTGGACGAAAGAGTCTGTAATTGATGACCGGCGCATTTGGAAATAC 111
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 15 LeuArgLysGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 34

QY 112 AATACATCCACATTTGGAAAGCATTAATATCACTGCTCCCAAGCTTAAGAACGAAAG 171
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 35 AlaAlaSerSerValArgGlySerLeuAlaValAlaSerValAlaValAlaValArg 54

QY 172 TTGCAACAGGACAAAGTGAATTACAGAGCTATCCAGCATTCAGGAAACATTAAGGTT 231
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 55 AlaArgGlyGlyAlaVal-----SerAlaArgTyrValLeu 66

QY 232 GACATTGATTGAGTCAG-----AGGTT 255
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 67 ProAspGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 86

QY 256 GTAGTTTACATCAAGCTCCCAAGATGTTGCTCTCTTCACAGCTTTTTCACAT 315
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 87 ValValLeuLeuSerGlnGlySerArgGlnGlyGlnLysLeuArgGlnGlnSerAlaAlaArg 106

QY 316 GTACTCTGGGTAATAGAGAGAGAGCTCAACTCT-----GTTCACTGCTGCA 366
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 107 ValValLeuThrSerLeuLeuAlaCysLeuProAlaGlyProArgValTyrPheLeuLys 126

QY 367 GGTGGGTTGCTAGTCTCTGCTGTTTCCCTGGCTCTGT----- 408
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 127 GlyGlyTyrGlnThrPheTyrSerGlnTyrProGlnCysCysValAlaPheValLysProIle 146

QY 409 -----GAAGAAATCACTAGTCTAGTCCCTACCTGCAATTTCTCAGCCTGC 453
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 147 SerGlnGlnLysIleGlnSerGlnArgAlaLeuIleSerGlnCysGlnLys----- 163

QY 454 TTACCTGTTGCCAAT-----GGGCAACCCCAAT 486
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

Db 164 ---ProValValaValSerTyrArgProAlaTyrArgProValGlnIle 182
QY 487 CTTCSSAATTTTATCTGGCTGSCAGAGATGCTCTCAACAGAGCTGATACAGAG 546
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 183 LeuProPheLeuLysLeuGlySerAlaTyrAlaAlaSerLysCysGlnPheLeuAla 202

QY 547 AATGGATGGTATGATGAATGCGAGCTATGATGCTCAAGCTGCTGATATCC 606
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 203 LeuIleIleThrAlaLeuLeuAlaValSerArgArgThrSerGlnAla---CysMetThr 221

QY 607 GAGTCTCATTTCTGCTGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTGCGCTG 666
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 222 HisLeuIleTyrLysTyrIleProValGlnArgSerHisThrAlaPheIleSerHis 241

QY 667 TTGGACAAATCATAGATTTCATTTGAGAAAGCAAAAGCTCCCAATGATGTTCTAGT 726
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 242 PheGlnGlnAlaIleAlaPheIleAlaPheCysValArgGlnLysGlyLysValLeuVal 261

QY 727 CAGCTTTAGCTGGAGATCTCCGCTCCGACCATGCTATGCGCTGATCATGAGAGG 786
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 262 HisCysGlnAlaGlyIleSerArgSerProThrIleCysMetAlaTyrLeuMetLysThr 281

QY 787 ATGACATGCTTTAGATGAGCTTACAGATTGTTGAAAGAAAGAAAGCTACTATATCT 846
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 282 LysGlnPheArgLysGlnAlaPheAlaPheArgTyrIleLysGlnArgArgSerMetValSer 301

QY 847 CCAACTTCATTTTCTGGGCAACTCTGCTGATGAGAAAGATTAAGAACAGACT 906
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 302 ProAsnPheGlyPheMetGlyGlnLeuGlnTyrLysSerGlnIleLeuProSerThr 321

RESULT 10
T46405
hypothetical protein DKFZp43401321.1 - human
C/Species: Homo sapiens (man)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 21-Jun-2002
R/Baum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wleemann, S.
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23034
A/Accession: T46405
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-303 <AAA>
A/Cross-references: EMBL:AL137704
A/Experimental source: adult testis; clone DKFZp43401321
A/Note: DKFZp43401321.1

Alignment Scores:
Pred. No.: 7.05e-21 Length: 303
Score: 325.00 Matches: 81
Percent Similarity: 52.65% Conservative: 38
Best Local Similarity: 35.84% Mismatches: 81
Query Match: 19.98% Indels: 26
Gaps: 6

US-10-029-345a-108_copy_538_1443 (1-906) x T46405 (1-303)

QY 266 ATCAAGCTCCCAAGATGTTGCTCTCTCTTCAAGACTGTTTCTACATGATCTGTG 325
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 29 ValArgAlaProSerMet-AlaLeuGlyValSerGlnLeu----- 41

QY 326 GTAACTGAGAAAGAGCTTCAACTCTGTTTCACTGCTGAGGTGGTGTGCTGAGTTCT 385
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 42 ---AlaGlyArgSerArgCysLeuCysSerGlnSer-GlnGlyGlyTyrGlnArgPhe 60

QY 386 CTCGTTGTTCCCTGCTGCTGCTGAGAAAGAAATTCACCTCA-----G 427
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 60 SerSerGlnTyrProGlnPheCysSerLysThrLysAlaLeuAlaAlaIleProGlyPro 80

QY 428 TTCCTACTGCAATTTCTCAGCT-----TGC-----TTACCTGTTG 463
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 80 AlaProProSerAlaThrGlnPheLeuAlaPheLeuGlnCysSerSerCysGlnThrProLeu 100

```


QY 464 CCAACATT---GGGCCAACCCGAATTTTCCCAATCTTATCTTGCCGTCAGGAGATG 520
 Db :|||
 QY 100 isapnglmglyglProvalgluileuProPheleuTyrlleuGlySerlathia 120
 QY 521 TCCTCAACAGAGCTGATACAGAGAAATGGGATGGTATGTTAAATGCCAGCTATA 580
 Db :|||
 QY 120 laalaargarapmetleuapalaaleuGlyllemraaleuLeuAsnValSerSera 140
 QY 581 CCTGTCCAAAGCCTGACTTATCCCGAGCTCATTTCTGCGTGTGCTGTGAATAGACA 640
 Db :|||
 QY 140 spCyPProasn---HisPhegluGlyhstYrglnTylysCysileProvalgluAspa 159
 QY 641 GCTTTTGGAGAAATTTTGGCCGTGGCAATATGACATGATGATTCATTTGAGAAACAA 700
 Db :|||
 QY 159 snHisLysAlaaspIleSerSerThrPheMetGluAlaileGlyTyrlleapalaVal 179
 QY 701 AAGCTCCCAATGATATGTTCTAGTGCATCTTTAGCTGGATCTCCGCTCCGACCA 760
 Db :|||
 QY 179 ysapCySargGlyhArgValleuValHisCysGlnAlaGlylleserArgserlathri 199
 QY 761 TCGCTATCGGCTTACATATGAGAGATGATGATGATCTTATGATGATGATGATGAT 820
 Db :|||
 QY 199 lEcysleuAlaTyrlleuMetMetLysLysArgValArgleuGluAlaPheGlnPhe 219
 QY 821 TGAAGAAAAGAGCTTACTATATCTCCAACTTTCTGAGGCAATCTCTGAGACT 880
 Db :|||
 QY 219 allysglnArgArgserIleIleSerProAsnPheSerPheMetGlyGlnleuGlnP 239
 QY 881 ATGAGAAAGAAAT 894
 Db :|||
 Db 239 heGlnserGlnVal 243

RESULT 11

T32494

hypothetical protein C05B10.1 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #ext_change 29-Oct-1999

R/Gelbel, C.; Wamsley, P.
 submitted to the EMBL Data Library, December 1997

A/Description: The sequence of C. elegans cosmid C05B10.

A/Reference number: 221178

A/Accession: T32494

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-365 <GBI>

A/Cross-references: EMBL:AF036685; PIDN:AMB8308.1; GSPDB:GN000622; CESP:C05B10.1

A/Experimental source: strain Bristol N2; clone C05B10

A/Genes: CESP:C05B10.1

A/Map position: 4

A/Intons: 22/2; 46/2; 74/3; 163/2; 204/2; 248/1; 287/2

Alignment Scores:

Pred. No.: 1,166-17
 Score: 289.00
 Percent Similarity: 45.79%
 Best Local Similarity: 24.92%
 Query Match: 17.76%
 DB: 2
 Gaps: 7

US-10-029-345A-108_COPY_538_1443 (1-906) x T32494 (1-365)

QY 19 GAACTCAATTTGTTACTGAGAGTTGTTGCTCTGCTGGAAGTGAAGCAAGAAAAGTG 78
 Db :|||
 QY 12 GylgluGulIleSerAlaGluGlnPheAsnArgIlePheHisGlu-----ArgAsnVal 29
 QY 79 CTGCTAATTATAGCCCGCATTTGTGGAATACATATACATCCACATTTTGAAGCCATT 138
 Db :|||
 QY 30 lIleValleuAspCysArgSer-----AmGlyAspSerValLysArgAlaAsn 45
 QY 139 AATATCAACTGCTCAAGCTTATGAAAGCAAGGTTGCAACAGCAACAAAGTTATATTACA 198

Db :|||
 QY 46 ArgleuArgleuProAlaIleuGlnArgIleuMetGlyGlySerMetArgleuSer 65
 QY 199 GAGCTCATCCAGCATTCACAGCAAGCAATAGGTGACATGATTTGCACTCAAGAGTTTGA 258
 Db :|||
 QY 66 ThrValProAspLeuLysAspLeuAsnSerProAspGlnCysProGlnValleuLeu 85
 QY 259 GTTTACATCAAGAGCTCCCAAGT-----GTTCGCTCTCTCT 297
 Db :|||
 QY 86 lIleProGlyAspSerGluGlnAspGluGlnLeuSerAlaAlaIleuAlaArgAsnLeu 305
 QY 298 TGAAGCTGTTTCTCACTGATCTTCTGGGTAACTGGAGAGAGCTTCACTCTGTTAC 357
 Db :|||
 QY 106 SerAsnHisTyrglnHisPheValleuGlyGlnProValLysThrleuLeuSerGlnPhe 125
 QY 358 CTGCTTGACAGGTGGTTGGCTGAGTTCTC-----GTTGTTTC 396
 Db :|||
 QY 126 ProThrleuArgAspAlaAlaAspGluAsnThrPheAsnThrPheGlnMetAsnSerMet 145
 QY 397 CTTGCTCTGTGAGAGAAATCCACTTATGCTTCCCTACCTGATTTCTCAGCTTCTTGA 456
 Db :|||
 QY 146 ProGlyGlnAlaSerGlyGlnGln-----AlaSerGlyProleuLeu 160
 QY 457 CTTGTCACCAACTTGGG----- 474
 Db :|||
 QY 161 AsnleuAsnGlnleuArgleuGlnGlyAspGlnGlyLysGlnArgAlaGluPhe 180
 QY 475 CCAACCCGAATCTTCCCAATTTTATCTTGGCTGCCAGAGATGTCCTCCCAAGAG 534
 Db :|||
 QY 181 ProValLysleuThrAsnPheleuTyrlleuGlyAsnAlaGluThrAlaLysAsnAsp 200
 QY 535 CTGATACAGACAGATGAGATGTTGTTATGTTATGTTATGTTATGTTATGTTATGTT 594
 Db :|||
 QY 201 ValLeuLysLysTyrlleuSerIleSerHisValIleAsnAlaThrAsnLeuProAsnThr 220
 QY 595 ---GACTTATCCCGAGCTCATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
 Db :|||
 QY 221 PheGlnGluAspProAsnMetArgTyrlleuArgIleSerAlaAspAsnAsnIleSerHis 240
 QY 652 AAAATTTTCCGCTGTTGGAACAATGATGATGATTTGATGAGAAAGCAAGCCCTCCAT 711
 Db :|||
 QY 241 AsnleuThrLysPhePheProGlnAlaIleSerPheIleAspAspAlaArgArgAsnAsp 260
 QY 712 GAGATGTTCTAGAGCACTGTTATGCTGGAATCTCCGCTGCCAGCATGCTATGCGCC 771
 Db :|||
 QY 261 SerAlaCysleuValHisCysleuAlaGlyIleSerArgSerValThrIleCysleuAla 280
 QY 772 TACATCATGAAGAGATGAGATGCTTTAGATGAAGCTTACATGTTTGAAGAAA 831
 Db :|||
 QY 281 TyrlleuMetLysThrGluMetCysThrleuAspSerAlaTyrglnThrValGlnLysArg 300
 QY 832 AGACTTACTATATCTCCAACTTTCTGCGCCCACTCTGATCTGATGAGAAAG 891
 Db :|||
 QY 301 AsnAlaSerIleAlaProAsnPheHisPheMetGlyGlnleuThrAspTyrglnLysMet 320
 QY 892 ATT 894
 Db :|||
 QY 321 lIleu 321

RESULT 12

T32698

Protein tyrosine phosphatase - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #ext_change 03-Dec-1999

C/Accession: T32698

R/Wood, V.; Skelton, J.; Churcher, C.M.; Rajadream, M.A.; Barrett, B.G.

submitted to the EMBL Data Library, July 1999

A/Reference number: 221870

A/Accession: T32698

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-330 <MOO>

A/Cross-references: EMBL:AL09652; PIDN:CAB51765.1; GSPDB:GN00067

A:Experimental source: strain 972h-; cosmid c17A3
 C:Genetics:
 A:Gene: p1041
 A:Map position: 2

Alignment Scores:
 Pred. No.: 8.01e-12 Length: 330
 Score: 223.50 Matches: 53
 Percent Similarity: 52.59% Conservative: 18
 Best Local Similarity: 39.26% Mismatch: 63
 Query Match: 13.74% Indels: 1
 DB: 2 Gaps: 1

US-10-029-345A-108_COPY_538_1443 (1-906) x T39698 (1-330)

QY 478 ACCCGAATTCCTCCCAATCTTTATCTTGGCTGCCAGCAGATGTCTTCACAGAGCTG 537
 Db 48 SerGluLeSerLyAsnLeuTyrLeSerSerTrpLySerGluLeuValSer 67

QY 538 ATACAGCAGATGGGATGGTATGTGTAAATGCCAGCTTACTCTGCCAAAGCTGAC 597
 Db 68 ThrSerApLySGlyLeaApLyThrLeuSerAlaMetSerIleAsnProAsnLeuSer 87

QY 598 TTATATCCCGAGCTCATTTCTGCGTGGCTGGATGACAGCTTTTGAGAAATT 657
 Db 88 ---ValProGluGlnGlnIleLeuTrpLeuGlnIleGluAspSerSerSerGlnAsnIle 106

QY 658 TTGCGCTGGTTGACAAATCAGTAGATTTCATTGAGAAAGCAAGCTCCCATGATGT 717
 Db 107 LeuGlnTyrPheGluLySerAsnLyPheIleAlaPheAlaLeuSerLyAsnAlaLySer 126

QY 718 GTTCTAGTGCAGCTGTTTAGCTGGGATCTCCGCTCCGCCACCATGCTATGCTTACATC 777
 Db 127 ValIeuValHisCyPheAlaGlyIleSerArgSerValThrIeuValAlaAlaLyTrpIeu 146

QY 778 ATGAAGAGGATGAGCATGCTTTAGATGAGCTTACAGATTGTGAAAGAAAGAACCT 837
 Db 147 MetLySerGluAsnMetTrpAsnThrGluGlnIleAlaLeuSerHisIleAsnGluArgTrpSer 166

QY 838 ACTATATCTCCAACTTCATATTTCTGGGCCCACTCTCGAGCAT 882
 Db 167 GlyIleSerProAsnAlaAsnPheLeuArgGlnIleuArgValTyr 181

RESULT 13
 T16056
 hypothetical protein F13D11.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001
 C:Accession: T16056
 R:Fullon, L.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid F13D11.
 A:Reference number: S69020
 A:Accession: T16056
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-186 <FNU>
 A:Cross-references: EMBL:U40939; NID:g1072175; PID:g1072176; PIND:AAA81700.1; CESP:F13D11.3
 A:Gene: CESP:F13D11.3
 A:introns: 30/3; 57/3; 85/3; 125/3; 172/3

Alignment Scores:
 Pred. No.: 9.46e-12 Length: 186
 Score: 222.50 Matches: 51
 Percent Similarity: 58.16% Conservative: 31
 Best Local Similarity: 36.17% Mismatch: 50
 Query Match: 13.68% Indels: 9
 DB: 2 Gaps: 3

US-10-029-345A-108_COPY_538_1443 (1-906) x T16056 (1-186)

QY 478 ACCCGAATTCCTCCCAATCTTTATCTTGGCTGCCAGCAGATGTCTTCACAGAGCTG 537
 Db 48 SerGluLeSerLyAsnLeuTyrLeSerSerTrpLySerGluLeuValSer 67

QY 538 ATACAGCAGATGGGATGGTATGTGTAAATGCCAGCTTACTCTGCCAAAGCTGAC 597
 Db 68 ThrSerApLySGlyLeaApLyThrLeuSerAlaMetSerIleAsnProAsnLeuSer 87

QY 598 TTATATCCCGAGCTCATTTCTGCGTGGCTGGATGACAGCTTTTGAGAAATT 657
 Db 88 ---ValProGluGlnGlnIleLeuTrpLeuGlnIleGluAspSerSerSerGlnAsnIle 106

QY 658 TTGCGCTGGTTGACAAATCAGTAGATTTCATTGAGAAAGCAAGCTCCCATGATGT 717
 Db 107 LeuGlnTyrPheGluLySerAsnLyPheIleAlaPheAlaLeuSerLyAsnAlaLySer 126

QY 718 GTTCTAGTGCAGCTGTTTAGCTGGGATCTCCGCTCCGCCACCATGCTATGCTTACATC 777
 Db 127 ValIeuValHisCyPheAlaGlyIleSerArgSerValThrIeuValAlaAlaLyTrpIeu 146

QY 778 ATGAAGAGGATGAGCATGCTTTAGATGAGCTTACAGATTGTGAAAGAAAGAACCT 837
 Db 147 MetLySerGluAsnMetTrpAsnThrGluGlnIleAlaLeuSerHisIleAsnGluArgTrpSer 166

QY 838 ACTATATCTCCAACTTCATATTTCTGGGCCCACTCTCGAGCAT 882
 Db 167 GlyIleSerProAsnAlaAsnPheLeuArgGlnIleuArgValTyr 181

[illegible]

[illegible]

```
QY      874 CTGCACATGAGAGAGATTAA 897
      ::||| ||| ||| :::::
Db      142 ValAspYrGInuYsArgLeuArg 149
```

RESULT 15

hypothetical protein C04F12.8 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*

C/Accession: T18915
#sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

Submitted to the EMBL Data Library, November 1996
A;Reference number: Z19044

A;Accession: T18915

A;status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A;Residues: 1-272 <WIL>

A; Cross-references: EMB
A; Experiments: 1

A/Experimental source: clone C04F12
C/Genetics: 357DB:GN00019; CESP:C04F12.8

A;Gene: CESP:C04F12.8

A:Map position: 1
A:Introns: 53/1 94/3 200/5

Alignment Scores:

Pred. No.:	1,91e-10	Length:	272
Score:	208.00	Matches:	49
Percent Similarity:	54.0%	Conservative:	31
Best Local Similarity:	33.11%	Mismatches:	62
Query Match:	12.78%	Indels:	6
	2	Cons:	6

95-10-029-345A-108_COPY_538_1443 (1-906) X T18915 (1-272)

Oy 466 AACATTGGG-----CCAAACCGAATCTTCCCAATCTTATCTTGCTGCTCCAG 513
 Db |||:|||||
 4 AasnValGlyValleuGlyGlnmetSerGlnIleAsnAspHisLeuValLeuSer--Gly 22
 Oy |||:|||||
 514 CGAGATGTCCTCCACAAAGAGCTGATACACGCAAGATGGATGTTGTTATGTTAAATGCC 573
 Db |||:|||||
 23 AlaGlyValleuLysProAspLysIleLysGlnAAGlySileAsnMetIleValAsnAla 42
 Oy AGCTATACCTGTCCAAAGCTGACTTTATCCCGGAGTGCATTTCTCGCTGTCCTGTG 633
 Db |||:|||||
 43 ThrThrGlnGluProSer---ThrIlyMetGlnGlyValaLapHrmIlySileAAGlyle 61
 Oy AATGACAGCTTTGTGAGAAATTTTCCGTCGCTTGGCAATCAAGTATGTTTCATTGAG 693
 Db |||:|||||
 62 GluAspHisProIlyrAlaArgleuAsnGlnHisPheAspValValaLapAspLysIleArg 81
 Oy AAGCAAAAGCCTCCCAATGATGTGTTCTAGAGCACTGTTTACTGCGATCTCCGCTCC 753
 Db |||:|||||
 82 AsnValIlyseGluArgGlyGlyLysThrIleuValHisCysMetAlaGlyValSerArgSer 101
 Oy GCACACATGGCATGCGCTCAATCATCATGAAGAGATGACATGCTCTTAAATCAACCTTAC 813
 Db |||:|||||
 102 AlaSerIleuValMetIleLysIleuValIlyHisGluHisIleMetThrIleuArgGlnAlaLys 121
 Oy AGATTGTGAAAGAAAAAGACCTACATATATCCCAAACCTCAATTTCTGGGCGCACTC 873
 Db |||:|||||
 122 HisIlyValIlyValaLalArgProIleLysArgProAsnValGlyPheThrIlyGlnMet 141

```
QY      874 CTGGACCTATGAGAAAGATTAC 897
Db      142 ValAsPTyRGInLysArgLeuArg 149

Search completed: June 21, 2004, 12:41:55
Job time : 21.7715 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:21 ; Search time 9.52534 Seconds
(without alignments)
9905.280 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 1627

Sequence: 1 atggccacatgagatgatgtg9.....agaagattaagaaccagact 906

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10029345/runat.21062004.122815.4177/bpp.query.fasta.1.3278
-DB=SwissProt.42 -QFMT=fastan -SUFFIX=erp -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USFR=US10029345.OCGN.1.1.33@runat.21062004.122815.4177 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt.42.:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1552	95.4	665	DUSG_HUMAN	Q9bya4 homo sapien
2	988	60.7	625	DUS8_HUMAN	Q13202 homo sapien
3	985	60.5	663	DUS8_MOUSE	O09112 mus musculu
4	469	28.8	482	DUSA_HUMAN	Q9y6w6 homo sapien
5	468	28.8	483	DUSA_MOUSE	Q9es80 mus musculu
6	467.5	28.7	381	DUS6_HUMAN	Q18828 homo sapien
7	462.5	28.4	381	DUS6_MOUSE	Q9dbd1 mus musculu
8	461.5	28.4	381	DUS6_RAT	Q64346 rattus norv
9	450	27.7	619	VHPI_CABEL	Q10038 caenorhabd
10	431	26.5	375	DUS4_CHICK	Q9pw71 gallus gall
11	427	26.2	394	DUS4_HUMAN	Q13115 homo sapien
12	425	26.1	367	DUS1_MOUSE	P38563 mus musculu
13	423.5	26.0	395	DUS2_RAT	Q62767 rattus norv
14	418	25.7	314	DUS2_HUMAN	Q05993 homo sapien
15	415	25.5	367	DUS1_HUMAN	P28562 homo sapien
16	415	25.5	367	DUS1_RAT	Q64623 rattus norv
17	409	25.1	318	DUS2_MOUSE	Q05922 mus musculu
18	408	25.1	384	DUS9_HUMAN	Q99956 homo sapien

19	385.5	23.7	320	DUS7_HUMAN	Q16829 homo sapien
20	382.5	23.5	320	DUS7_MOUSE	Q91246 mus musculu
21	358.5	22.0	384	DUS7_RAT	O54898 rattus norv
22	357.5	22.0	280	DUS7_RAT	O63340 rattus norv
23	351	21.6	384	DUS5_HUMAN	Q16690 homo sapien
24	262	16.1	198	DUS6_HUMAN	Q95147 homo sapien
25	261	16.0	198	DUS6_MOUSE	Q9j147 mus musculu
26	205.5	12.6	188	DUS1_HUMAN	O8ne10 homo sapien
27	205.5	12.6	217	DUS1_MOUSE	O8wt22 homo sapien
28	205	12.6	223	STYX_MOUSE	O60959 mus musculu
29	204	12.5	223	STYX_HUMAN	O8wu10 homo sapien
30	202	12.4	364	PVH1_YEAST	Q02256 saccharomyc
31	201.5	12.4	313	MSY1_HUMAN	Q9y618 homo sapien
32	195.5	12.0	185	DUS3_HUMAN	P51452 homo sapien
33	194	11.9	185	DUS3_MOUSE	O9d743 mus musculu
34	185	11.4	807	YB9T_YEAST	P38148 saccharomyc
35	183.5	11.3	276	PTP3_CHLEU	Q39491 chlamydomon
36	181.5	11.2	278	PMP1_SCHPO	O13453 schizosacch
37	178.5	11.0	295	DUSF_HUMAN	O9h122 homo sapien
38	177	10.9	125	DUSF_MOUSE	O8r442 mus musculu
39	171.5	10.5	339	DUSC_MOUSE	O9d052 mus musculu
40	171	10.5	469	MSG5_YEAST	P38580 saccharomyc
41	158.5	9.7	340	DUSC_HUMAN	O9un16 homo sapien
42	157	9.6	209	YIL3_YEAST	P40479 saccharomyc
43	155	9.5	198	DUSD_HUMAN	Q9u116 homo sapien
44	152.5	9.4	597	PTPX_CANAL	P43078 candida alb
45	152	9.3	292	Y042_CABEL	P34680 caenorhabd

ALIGNMENTS

RESULT 1
DUSG_HUMAN
ID DUSG_HUMAN STANDARD; PRT; 665 AA.
AC Q9bya4; Q9COG3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein kinase phosphatase 16 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 7) (MAP kinase
phosphatase 7) (MKP-7).
DE DUSP16 OR MKP7 OR KIAA1700.
GN DUSP16
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21486429; PubMed=11489891;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.,
RT "MKP-7, a novel mitogen-activated protein kinase phosphatase,
RT functions as a shuttle protein.";
RL J. Biol. Chem. 276:39002-39011 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.,
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:347-355 (2000).
CC - FUNCTION: Involved in the inactivation of MAP kinases.
CC - CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein +
CC tyrosine + phosphate.
CC - CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC - SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC - SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC - SIMILARITY: Non-receptor class dual specificity subfamily.
CC - SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration


```

CC      This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
CC      EMBL: U27193; AAA83151.1; -.
CC      DR HSSP; Q16828; IMKP.
CC      DR GENEW; HGNC:3074; DUSP8.
CC      MIM; 602038; -.
CC      DR GO; GO:0005737; C:cytoplasm; TAS.
CC      DR GO; GO:0004723; F:protein tyrosine phosphatase activity; TAS.
CC      DR GO; GO:0000188; P:inactivation of MAPK; TAS.
CC      DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
CC      DR InterPro; IPR000340; DS phosphatase.
CC      DR InterPro; IPR008343; MAPK phospho.
CC      DR InterPro; IPR001763; Rhodanese-like.
CC      DR InterPro; IPR000387; TYR_phosphatase.
CC      DR Pfam; PF00782; DSPC; 1.
CC      DR PRINTS; PR01764; MAPKPHPTASE.
CC      DR SMART; SM00195; DSPC; 1.
CC      DR SMART; SM00450; RHOD; 1.
CC      DR PROSITE; PS50206; RHODANES_3; 1.
CC      DR PROSITE; PS00383; TYR_PHOSPHATSE_1; 1.
CC      DR PROSITE; PS50056; TYR_PHOSPHATSE_2; 1.
CC      DR PROSITE; PS50054; TYR_PHOSPHATSE_DUAL; 1.
CC      KW Hydrolase; Nuclear protein.
CC      FT DOMAIN 23 138 RHODANES.
CC      FT DOMAIN 162 430 PROTEIN-TYROSINE PHOSPHATASE.
CC      FT DOMAIN 310 550 PRO-ICH.
CC      FT ACT_SITE 246 246 PHOSPHOCSTEINE INTERMEDIATE (BY
CC      FT SMILARITY).
CC      SQ SEQUENCE 625 AA; 65840 MW; DCEBA14487219666 CRC64;
CC
CC      Alignment Scores:
CC      Pred. No.: 2,61e-84 Length: 625
CC      Score: 988.00 Matches: 190
CC      Percent Similarity: 81.40% Conservative: 55
CC      Best Local Similarity: 63.12% Mismatches: 54
CC      Query Match: 60.73% Indels: 2
CC      DB: 1 Gaps: 2
CC
CC      US-10-029-345A-108_COPY_538_1443 (1-906) x DUS8_HUMAN (1-625)
CC
CC      QY 1 ATGGCCCATGAGATGATTCGAACTCAATTCGTT---ACTGAGAGTTGGTGCCTGCTGCT 57
CC      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC      Db 1 MetAlaGlyAspArgLeuProArgIyValMetAspAlaLysLeuAlaSerLeu 20
CC      58 GAAATGTGAGACGAAAGAGTCGTCGAATTGATTAGCGCGCATTTGTGGATATCAATACA 117
CC      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC      Db 21 ArgIyGlyProGlyIyLProLeuValIleAspSerIyGserPheValGluIyIyAsnSer 40
CC      118 TCCCAACATTTTGGAGACCATTAATATCAATGCTCCCAAGCTTATGAAGGAGATTGCAA 177
CC      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC      Db 41 TrpHisValIeuSerSerValaenIleCysCySerIyLysLeuValIyArgIyGluIn 60
CC      178 CAGACCAAGAAGTGTAAATACAGAGCTCATCCGCAATTCAGCGAAACATRAAGTTGACATT 237
CC      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC      Db 61 GInGlyLysValaThrIleAlaGluLeuIleGInProAlaIlaIyArgSerGInValGluAla 80
CC      238 GATTCCAGTCAGAGAGTGTTAGTTATTCAGATCAAAAGCTCCCAAGATGTTGCCCTCTCTCT 297
CC      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC      Db 81 ThrGluProGInAspValaValaIyIyArgGInSerIyIyArgAspAlaSerValaLeuAla 100
CC      298 TCAGACTGTTTTCATCATGTACTTCTGGGTAAATCTGAGAGAGCTTCAACTGTGTCAC 357
CC      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC      Db 101 AlaAspSerPheLeuSerIleLeuLeuSerIyLysLeuAspGlyCysPheAspSerValaAla 120
CC      358 CTGCTTCAGAGGTGGTTCGTAAGTTCTCTGTTTTCCTGGGCTCTGTGAAGAGAAA 417
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

[illegible]

CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC

 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR	EMBL; X95518; CAA64772.1; ..
DR	HSSP; Q16828; IMKP.
DR	MGD; MG1:106626; Disps.
DR	InterPro; IPR000340; DS_phosphatase.
DR	InterPro; IPR008343; MAPK_phosph.
DR	InterPro; IPR001763; Rhodanese-like.
DR	InterPro; IPR000387; TYR_phosphatase.
DR	Pfam; PF00782; DSPC; 1.
DR	PRINTS; PR01764; MAPKPHPTASE.
DR	SMART; SMO0195; DSPC; 1.
DR	SMART; SMO0450; RHOD; 1.
DR	PROSITE; PS50206; RHDNANESE_3; 1.
DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR	PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
KW	Hydrolase; Nucleic Protein.
FT	DOMAIN 23 138
FT	DOMAIN 162 432
FT	DOMAIN 452 459
FT	DOMAIN 555 558
FT	DOMAIN 559 576
FT	DOMAIN 577 600
FT	DOMAIN 311 552
FT	ACT_SITE 246 246
SEQUENCE	663 AA; 68847 MW; 4166429A1C1FA7C CRC64;
SO	SIMILARITY)

Fragment Scores:	
Pred. No.:	5.03e-84
Score:	985.00
Percent Similarity:	80.73%
Best Local Similarity:	62.46%
Query Match:	60.54%
DB:	1
Length:	663
Matches:	188
Conservative:	55
Mismatches:	56
Indels:	2
Gaps:	2

US-10-029-345A-108_COPY_538_1443 (1-906) x DUS8_MOUSE (1-663)

```

1 ATGGCCCATGAGCTGATTTGGAACTCAAATTGTT---ACTGAGAGGTTGGCTGGCTGGT
2      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
3 1 MetalaglyAspharygLeuprArygLyAspValMetaspAlaySlySLeuAlaSerLeuLeu
4      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
5 58 GAAAGTGGAAACGGAAAAAGTGGCTGCTTAATTATAGCCGGCCCATTTGGTGAATCAATCA
6      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7 21 ArgglygLygProglygLyproLeuValTllempserhgerSerhethalGlnutyrenSer
8      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
9 118 TCCCAACATTTTGGAAAGCCATTAATATCACTGCCTCCCAAGTTATGAAGGAAAGTTGCA
10      ||||| ||||| ||||| ||||| ||||| ||||| |||||
11 41 CysHsValIleuSerSerValSnnIleCysCysSerLySLeuValLySArgyLeuGln
12      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
13 178 CAGGCAAAAGGTTAATTACAGAGCTTCATCCAGCATTCAGCAAAACATPARGTGGACATT
14      ||||| ||||| ||||| ||||| ||||| ||||| |||||
15 61 GlnglyySValTrIlealagIleuIleGlnProAlaThrArgSerClnValSspAla
16      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
17 238 GATTCCAGTCAGAAAGTTGTAGTTATACGATCAAAAGTCCCAAGATGTTCCCTCTCTCT
18      ||||| ||||| ||||| ||||| ||||| ||||| |||||
19 81 ThrGlnuProGlnAspValValValTlyrAspGlnSerTrnArgAspAlaSerValIleuAla
20      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
21 298 TCAGACTGTTTTCATCTACTGACTTCCTGGGTAACTGGAAGAAAGATTCATCTGTGTCAC
22      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
23 101 AlaAspSerhethuSerIleLeuLeuSerLySLeuAspGlyCysPheAspSerValAla
24      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
25 358 CTGCTTCAGAGTGGGTTGTCTGAGTTCCTCTGTTGTTCCCTGGGCCCTGTGAAGGAAA
26      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

```

[illegible]

RESULT 4

ID	DUSA_HUMAN	STANDARD,	PRT,	482 AA.
AC	O9Y6M6			
DT	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	10-OCT-2003	(Rel. 42, Last annotation update)		
DE	Dual specificity protein phosphatase 10 (EC 3.1.3.48)			
DE	(Mitogen-activated protein kinase phosphatase 5) (MAP kinase phosphatase 5) (MKP-5).			
GN	DUSP10 OR MKP5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99321929; PubMed=10391943;			
RA	Tanoue T., Moriyuchi T., Nishida E.;			
RT	"Molecular cloning and characterization of a novel dual specificity			
RL	phosphatase, MKP-5.";			
RN	J Biol. Chem. 274:19949-19956(1999).			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20065165; PubMed=10597297;			
RA	Theodosiou A., Smith A., Gilleron C., Arkinstall S., Ashworth A.;			
RT	"MKP5, a new member of the MAP kinase phosphatase family, which			
RT	selectively dephosphorylates stress-activated kinases.";			
RL	Oncogene 18:6981-6988(1999).			
RP	SEQUENCE FROM N.A.			
RX	TISSUE=Brain, Lung, and Testis;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse J.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			

Accession	Protein Name	Length	Score	Similarity
RA A00001	Alcalase	482	8.16e-36	57.72%
RA A00002	Hopkins	469	4.69e-00	57.72%
RA A00003	Diachenko	469	4.69e-00	57.72%
RA A00004	Scapleton	469	4.69e-00	57.72%
RA A00005	Brownstein	469	4.69e-00	57.72%
RA A00006	Rah S.S.	469	4.69e-00	57.72%
RA A00007	Boak S.A.	469	4.69e-00	57.72%
RA A00008	Richards S.	469	4.69e-00	57.72%
RA A00009	Vallalon D.K.	469	4.69e-00	57.72%
RA A00010	Faney J.	469	4.69e-00	57.72%
RA A00011	Whiting M.	469	4.69e-00	57.72%
RA A00012	Blakesley R.W.	469	4.69e-00	57.72%
RA A00013	Rodriguez A.C.	469	4.69e-00	57.72%
RA A00014	Butterfield J.S.N.	469	4.69e-00	57.72%
RA A00015	Schmerch A.	469	4.69e-00	57.72%
RA A00016	Generation and initial analysis	469	4.69e-00	57.72%
RT RT	human and mouse cDNA sequences	469	4.69e-00	57.72%
CC CC	-1- FUNCTION: Involved in the inactivation of MAP kinases. Has a specificity for the MAPK1/MAPK2/MAPK3/MAPK4 subfamily.	469	4.69e-00	57.72%
CC CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein tyrosine + phosphate.	469	4.69e-00	57.72%
CC CC	-1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein + phosphate.	469	4.69e-00	57.72%
CC CC	-1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.	469	4.69e-00	57.72%
CC CC	-1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family. Non-receptor class dual specificity subfamily.	469	4.69e-00	57.72%
CC CC	-1- SIMILARITY: Contains 1 rhodanese domain.	469	4.69e-00	57.72%
CC CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/).	469	4.69e-00	57.72%
CC CC	or send an email to license@isb-sib.ch	469	4.69e-00	57.72%
DR DR	EMBL; AB026436; BAB1668.1; -	469	4.69e-00	57.72%
DR DR	EMBL; AF179212; AAD51857.1; -	469	4.69e-00	57.72%
DR DR	EMBL; BC031405; AAH31405.1; -	469	4.69e-00	57.72%
DR DR	HSSP; Q16828; IMKP.	469	4.69e-00	57.72%
DR DR	Genew; HGNC:13065; DUSP10.	469	4.69e-00	57.72%
DR DR	GO; GO:0005277; C:cytoplasm; TAS.	469	4.69e-00	57.72%
DR DR	GO; GO:0005634; C:nucleus; TAS.	469	4.69e-00	57.72%
DR DR	GO; GO:0004721; F:protein phosphatase activity; TAS.	469	4.69e-00	57.72%
DR DR	GO; GO:0007254; P:JNK cascade; TAS.	469	4.69e-00	57.72%
DR DR	GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.	469	4.69e-00	57.72%
DR DR	GO; GO:0006950; P:response to stress; TAS.	469	4.69e-00	57.72%
DR DR	InterPro; IPR000340; DS_phosphatase.	469	4.69e-00	57.72%
DR DR	InterPro; IPR008343; MAPK_phosph.	469	4.69e-00	57.72%
DR DR	InterPro; IPR001763; Rhodanese-like.	469	4.69e-00	57.72%
DR DR	InterPro; IPR000387; TYR_phosphatase.	469	4.69e-00	57.72%
DR DR	Pfam; PF00783; DSPC; 1.	469	4.69e-00	57.72%
DR DR	Pfam; PF00581; Rhodanese; 1.	469	4.69e-00	57.72%
DR DR	PRINTS; PR01764; MAPKPHPTASE.	469	4.69e-00	57.72%
DR DR	SMART; SM00195; DSPC; 1.	469	4.69e-00	57.72%
DR DR	SMART; SM00450; RHOD; 1.	469	4.69e-00	57.72%
DR DR	PROSITE; PS50206; RHODANES_3; 1.	469	4.69e-00	57.72%
DR DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.	469	4.69e-00	57.72%
DR DR	PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.	469	4.69e-00	57.72%
DR DR	PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.	469	4.69e-00	57.72%
KW KW	Hydroxylase; Nuclear protein.	469	4.69e-00	57.72%
FT FT	DOMAIN 168 285 RHODANES.	469	4.69e-00	57.72%
FT FT	DOMAIN 384 453 PROTEIN-TYROSINE PHOSPHATASE.	469	4.69e-00	57.72%
FT FT	ACT_SITE 408 408 PHOSPHOTYROSINE INTERMEDIATE (BY SIMILARITY)	469	4.69e-00	57.72%
SQ SQ	SEQUENCE 482 AA; 52642 MW; A8CB74ABF9498CD4 CRC64;	469	4.69e-00	57.72%

Query Match:	28..83%	Indels:	30
DB:	1	Gaps:	7
US-10-029-345A-108_COPY_538_1443 (1-906) x DUSA_HUMAN (1-482)			
QY	CTGCTAATTGATAGCCGGCCATTGTGGAAATACATATCATCCCATTTTGGAAACATT	138	
Db	173 ValIleIleAerCybArGpProPheMetGluTyLAsnLySerHisIleGInGlyAlaVal	192	
QY	139 AATATCACTGCTCC---AAGCTTATGAAACGAAAGTTGGACACGACAAAGTTAATT	195	
Db	193 HisIleAsnCybAlaAerLySIIeSerArGArGInGInGInGlySIIeThrVal	212	
QY	196 ACAGAGCTCATCCGAC---CATTCAGCGAAACATTAAGTTGATGATGATGACATGACAG	252	
Db	213 LeuAerPheIIeSerCybArGInGlyLyAAsPseRheLybArGIIePheSerLySgIIu	232	
QY	253 GTTGCTAGTTACGATCAAGCTCCCAAGATGTGGCTCTCTCTTCAGACTGTTTCTC	312	
Db	233 IleIleValIyArbArGInAenThInAenGInPProSerArGAlMetProSerGInProLeu	252	
QY	313 ACTGTACTTCTGGTAAACTGAGACAGAGCTTCAACTCTGTTCACCTGCTTGACAGTGGG	372	
Db	253 HisIleValIleuGInSerIleuLySArGInGlyLySgIIuProLeuValIeulySgIIyGly	272	
QY	373 TTTCCTAGTCTCTCTGTTTTCCTCCCGCTCGTGAAGAA-----	417	
Db	273 LeuSerSerPheLySgInAenHisIleAenIleuCybArbAenSerIleuGInIleuGInGly	292	
QY	418 -----TCCACTGTAGCTCCCTACCTGACTTCT	444	
Db	293 CybArGInGlyAlGlyGlyAlaSerAlaIaIaSerSerIleuPProGInProIle---	311	
QY	445 CAGCTTGCTTACTCT---GTTGCCACATTGGGCGCAACCGAATCTTCCCAATCTTAT	501	
Db	312 ---ProThrThrProAerIIeGInAenAlAGInIleuThrProIleIleuProPheIleuPhe	330	
QY	502 CTTCGGCTCCGACGAGATGTCTTCACACAGAGCTGATACAGACGAATGGGATTCGTAT	561	
Db	331 LeuGInLySInGInAerAlaGInAerPLeuAerPThrMetGInIyGLeuAenIIeGlyTyR	350	
QY	562 GTGTTAATATGCAGAC-----TATACCTGTCCAAAGCTGACTTATATCC	606	
Db	351 ValIleAenValIThrThrHisIleuPProeulyThHisIyGInGlySgIleuPhe-----	368	
QY	607 GAGTCTCATTTCTCGCTGCTGCTGCTGATGACACCTTTTGTGAGAAATTTGGCCGTGG	666	
Db	369 -----AenTyRlySArGLeuPProAlaThArbSerAenHisIleAenIleuAerGInTyR	386	
QY	667 TTGACACAAATCAGATATTCATTGAGAAAGCAAAAGCTTCGAATGATGTGTTTATGTG	726	
Db	387 PheGInGInIuAlaPheGInIuPheIIeGInGInIuAlaHisIleGInCySgIySgIyLeuIleIle	406	
QY	727 CACGTTAGTGGGATCTCCCGTCCGACCACTGTAATGCGCTGATCAGTAAGAGG	786	
Db	407 HisIeCySgInIaGlyValIeArGSerAlaThrIleAlIleAlaTyRLeuMetLyHis	426	
QY	787 ATGACATGTCTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTACTATATCT	846	
Db	427 ThrArgMetThrMetThrAerAlaTyLybPheValIySgIySArGProIleIIeIleSer	446	
QY	847 CCAAACTTCAATTTTCTGGGCGCACTCTGACTATGAGAAGATTAAGAAC	900	
Db	447 ProAenIleuAenPheMetGInIleuIleuGInPheGInGInIuAerPLeuAenAen	464	
RESULT 5			
DUSA_MOUSE	STANDARD;	PRT;	483 AA.
AC	Q9ES50; Q9CZV9;		
DT	10-OCT-2003 (Rel. 42, Created)		
DT	10-OCT-2003 (Rel. 42, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)		

Db 408 HieGcglNaagIValserArgSerAlaThrIleValIleAlaTyrlaMetlySHis 447
 QY 787 ATGACATCTCTTAACTTACAGCTTACAGATTGTGAAAGAAAAAGACTACTATTC 846
 Db 428 ThrArgMetThrMetThrAspAlaTyrlaPheValIleGlySAlaGProIleIleSer 447
 QY 847 CCAACTCTCAATTTCTGGGCGCAACTCTGACGTACAGAGAAAGATTAGAAC 900
 Db 448 ProhenleuhenPheMetGlyGlnIleuGlnPheGlnIleuGlnPheleuhen 465
 RESULT 6
 DUS6_HUMAN STANDARD; PRT; 361 AA.
 ID DUS6_HUMAN
 AC 016826; 075109; Q9BSH6;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
 DE phosphatase 3) (MKP-3) (dual specificity protein phosphatase PYST1).
 GN DUSP6 OR MKP3 OR PYST1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Forebrain;
 RX MEDLINE=96312959; PubMed=8670865;
 RA Groom L.A., Sneddon A.A., Alessi D.R., Dowd S., Keyse S.M.;
 RT "Differential regulation of the MAP, SAP and RK/p38 kinases by Pyct1,
 RT a novel cytosolic dual-specificity phosphatase.";
 RL EMBO J. 15:3621-3632(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Liver;
 RX MEDLINE=99077745; PubMed=9858808;
 RA Furukawa T., Yatsuoka T., Yousef E.M., Abe T., Yokoyama T.,
 RA Fukushima S., Soeda E., Hoshi M., Hayashi Y., Sunamura M., Kobari M.,
 RA Horii A.;
 RT "Genomic analysis of DUSP6, a dual specificity MAP kinase phosphatase,
 RT in pancreatic cancer.";
 RL Cytogenet. Cell Genet. 82:156-159(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT VAL-114.
 RC TISSUE=Colon, Kidney, Skin, and Stomach;
 RX MEDLINE=2388257; PubMed=12477932;
 RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,
 RA Klausner R.D., Collins F.S., Moore T., Max S.I., Wang J., Heien F.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
 RA Dlatchenko L., Marutika K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchan J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywiński M.I., Skalska U., Smilins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 204-347.
 RX MEDLINE=99140299; PubMed=10048930;
 RA Stewart A.B., Dowd S., Keyse S.M., McDonald N.O.;
 RT "Crystal structure of the MAPK phosphatase Pyct1 catalytic domain and
 RT implications for regulated activation.";

RL Nat. Struct. Biol. 6:174-181(1999).
 CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
 CC family.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing, Named isoforms=2;
 CC Name=1;
 CC IsoId=Q16828-1; Sequence=Displayed;
 CC Name=2; Synonyms=DUSP6-ALT;
 CC IsoId=Q16828-2; Sequence=VSP_005137;
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the European Bioinformatics Institute and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL, X93920; CAA63813.1; -;
 CC EMBL, AB013601; BAA31968.1; -;
 CC EMBL, AB013383; BAA31968.1; JOINED.
 CC EMBL, AB013600; BAA31968.1; JOINED.
 CC EMBL, AB013382; BAA34369.1; -;
 CC EMBL, AB013602; BAA31969.1; -;
 CC EMBL, BC003143; AAH03143.1; -;
 CC EMBL, BC003562; AAH03562.1; -;
 CC EMBL, BC005047; AAH05047.1; -;
 CC EMBL, BC037236; AAH37236.1; -;
 CC PDB, 1MKP; 22-JUL-99.
 CC PDB, 1HZM; 25-JAN-02.
 CC GeneW; HGNC:3072; DUSP6.
 CC MIM; 602748; -;
 CC GO; GO:0005737; Cytoplasm; IDA.
 CC GO; GO:0005625; C:soluble fraction; TAS.
 CC GO; GO:0004722; Protein serine/threonine phosphatase activity; IDA.
 CC GO; GO:0004725; Protein tyrosine phosphatase activity; IDA.
 CC GO; GO:000188; P:inactivation of MAPK; IDA.
 CC GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
 CC InterPro; IPR000340; DS phosphatase.
 CC InterPro; IPR008143; MAPK phosph.
 CC InterPro; IPR001763; Rhodanese-like.
 CC InterPro; IPR000387; TYR phosphatase.
 CC Pfam; PF00782; DSpC; 1.
 CC Pfam; PF00581; Rhodanese; 1.
 CC PRINTS; PR01764; MAPKPHPTASE.
 CC SMART; SM00195; DSpC; 1.
 CC SMART; SM00450; RHOD; 1.
 CC PROSITE; PS50206; RHODANSE 3; 1.
 CC PROSITE; PS00383; TYR PHOSPHATASE 1; FALSE_NEG.
 CC PROSITE; PS50056; TYR PHOSPHATASE 2; 1.
 CC PROSITE; PS50056; TYR PHOSPHATASE DUAL; 1.
 CC HydroLase; Polymorphism; Alternative splicing; 3D-structure.
 FT DOMAIN 30 148 RHODANSE.
 FT 206 381 PROTEIN-TYROSINE PHOSPHATASE.
 FT 293 293 PHOSPHOCYSTEINE INTERMEDIATE.
 FT ACT SITE 134 279 Missing (in isoform 2).
 FT VARSP LIC 114 114 /FTId=VSP_005137.
 FT L -> V.
 FT FTId=VAR_015113.
 FT STRAND 208 211
 FT TURN 212 213
 FT STRAND 214 217
 FT TURN 219 220
 FT HELIX 225 230
 FT TURN 231 232

FT STRAND 233 238
 FT STRAND 246 250
 FT TURN 251 252
 FT STRAND 253 257
 FT TURN 263 265
 FT HELIX 264 271
 FT HELIX 269 272
 FT TURN 272 284
 FT STRAND 285 286
 FT STRAND 288 292
 FT HELIX 298 312
 FT TURN 313 313
 FT HELIX 316 326
 FT TURN 328 329
 FT TURN 335 336
 FT HELIX 337 345
 FT TURN 346 346
 SQ SEQUENCE 381 AA; 42333 MM; 03BC12252CE73B26 CRC64;

Alignment Scores:
 Pred. No.: 1,08e-35 Length: 381
 Score: 467.50 Matches: 110
 Percent Similarity: 53.56% Conservative: 63
 Best Local Similarity: 34.06% Mismatches: 105
 Query Match: 28.73% Indels: 45
 Gaps: 7

US-10-029-345A-108_COPY_538_1443 (1-906) x DUS6_HUMAN (1-381)

QY 55 CTGGAAGAGTGAACCGGAAAGTCTGCTTAATTGATACCGGCATTTGTGAATACAT 114
 Db 27 LeuGluLeuGlyAsnGluLeuLeuLeuMetAspCysArgProGlnGluLeuYrGlu 46
 QY 115 ACATCCCATATTTGGAAGCCATTAATCAACGCTCCCAAGCTTATGAAGCAAGTTG 174
 Db 47 SerSerHisTleGluSerAlaIleAsnValAlaIleProGlyIleMetLeuArgLeu 66
 QY 175 CAACAGACAAAGGTGTAATACAGAGCTCATCCAGATTCAAGC--AAACATAAGTT 231
 Db 67 GlnYsGlyAsnLeuProValArgAlaLeuPheThrArgGlyGluAspArgSerPhe 86
 QY 232 GACATGATTCG--AGTCAGAAAGTTGTAGTTAGATCAAGCTCCCAAGAT--GTT 285
 Db 87 ThrArgArgCysGlyThrAspThrValValLeuYrAspGluSerSerSerAspTrpAsn 106
 QY 286 GCTCTCTCTCTTCAAGCTGTTTTCACGTCGTAATCTTGGTAACTGAGAAGAGCTTC 345
 Db 107 GlnAsnThrGlyGlyGluSerLeuLeuGlyLeuLeuLeuLeuYrGlyLeuYrAspGluGly 126
 QY 346 AACTCTTCACTGCTGTCAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 405
 Db 127 CysArgGlnPheThrLeuGluGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 146
 QY 406 TGTGAAGAAATTCACCTAGTCCCTACCTCC--ATTCTCAGCCTTGTACCTGTT 462
 Db 147 CysGlu-----ThrAsnLeuAspGlySerSerSerSerSerProLeuProVal 164
 QY 463 GCCAATTTGGG----- 474
 Db 165 LeuGlyLeuGlyGlyLeuArgIleSerSerAspSerSerSerAspIleGluSerAspLeu 184
 QY 474 ----- 474
 Db 185 AspArgAspProAsnSerAlaThrAspSerAspGlySerProLeuSerAsnSerGlnPro 204
 QY 475 -----CCAACCGCAATTTCCCAATCTTTATCTTGGTGCAGAGAGATTCCTCAAC 528
 Db 205 SerPheProValGluIleLeuProPheLeuYrLeuGlyCysAlaIlyAspSerThrAsn 224
 QY 529 AAGAGCTGATACAGCAGATGGATTTGTTAGTTAAAGCCAGCTATACCTGTCCA 588
 Db 225 LeuAspValLeuGlnGluPheGlyIleYrYrIleLeuAsnValThrProAsnLeuPro 244
 QY 589 AAGCCT---GACTTATCCCGCAGGTCAATTTCTGCGTGTGCTGAGTAATGACAGCTTT 645

Db 245 AsnLeuPheGluAsnAlaGlyGluPheYrYrYsGlnIleProIleSerAspHisTrp 264
 QY 646 TGTGAGAAATTTTGGCGGTGGAGCAATACAGATTTCAATTTGAGAAAGAAAGCC 705
 Db 265 SerGlnAsnLeuSerGlnPhePheProGluAlaIleSerPheIleAspGlnAlaArgGly 284
 QY 706 TCCATGATGATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 765
 Db 285 LysAsnCysGlyValLeuValHisCysLeuAlaGlyIleSerArgSerValThrValThr 304
 QY 766 ATCCCTTCATCATGAGAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 825
 Db 305 ValAlaYrLeuMetGlnYsLeuAsnLeuSerMetAsnAspAlaYrAspIleValYs 324
 QY 826 GAAAAAGACCTACTCATATCTCCAAATTTCTGAGGCAACTCCTGAGATGAG 885
 Db 325 MetLysLysSerAsnIleSerProAsnPheAsnPheMetGlyGlnLeuLeuAspPheGlu 344
 QY 886 AAGAAAGATT 894
 Db 345 ArgThrLeu 347

RESULT 7
 DUS6_MOUSE STANDARD; PRT; 381 AA.
 AC Q9DBB1; Q9D7L4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
 DE phosphatase 3) (MKP-3).
 GN DUSP6 OR MKP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Liver, and Tongue;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirni L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gusticich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbarts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata C., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22389257; PubMed=12477932;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Altschuler R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Hopkins R.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bat N.K.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.P., Schaefer C.E.,
 RA Brownstein M.J., Urdin T.B., Toshiyuki S., Carninci P., Prange C.T.E.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

QY	555	CTGGAAATGGACAGCAAAAAAGTGGTCTGAATTGATAGCCGGCATTTGTGTGAATACAAAT	114
Db	27	LeuGluLeuGluGluGluGluGluGluLeuLeuLeuMetAspCysArgProGluGluLeuArgLu	46
QY	115	ACATCCCAATTTTGGAAAGCCATTATATCACTGTCTCAAGCTTATATGAAAGCAAGTTG	174
Db	47	SerSerHisIleGluSerAlaIleLeuSerValAlaIleProGluTyrIleMetLeuArgArgLeu	66
QY	175	CAACAGCAAAAGTGTAAATTACAGACTCAATCCAGCAAT--TCAGCAAACTAAGTT	231
Db	67	GluIleSerIAsnLeuProValArgAlaLeuSerHisArgCysGluAspArgAspArgIle	86
QY	232	GACATTGATTGC---AGTCAGAAAGTTGATTGATTAGATCAAAAGCTCCCAAGAT--GTT	285
Db	87	ThrArgArgCysGluTyrAspArgThrValValLeuTyrAspIleuSerSerAspArgAsn	106
QY	286	GCCTCTCTCTTCAAGACGCTGTTTCTCACTGTACTTGTGGTAAACTGCAAGAAAGCTTC	345
Db	107	GluAsnThrGlyGlyGlyGlySerValLeuGlyIleuLeuMetLeuValLeuLeuAspArgLu	126
QY	346	AACCTGTTCACTGCTTGGACAGTGGATTGCTGAGATTGTCGTTGTTTCCCTGGCTC	405
Db	127	CysArgAlaPheTyrLeuGluGluGlyGlyPheSerTyrPheGluAlaGluPheAlaLeuHis	146
QY	406	TCGTGAAGAAATACACACTAGTCCCTACCTGC--ATTCTCAAGCTTGGCTTACTGTT	462
Db	147	CysGlu-----ThrAsnLeuAspGlySerCysSerSerSerSerProProLeuProVal	164
QY	463	GCCAAACATTGGG-----	474
Db	165	LeuGlyLeuGlyGlyLeuArgIleSerSerAspSerSerSerSerAspIleGluSerAspLeu	184
QY	474	-----	474
Db	185	AspArgAspProAsnSerAlaThrAspSerAspGlySerProLeuSerAsnSerGlnPro	204
QY	475	-----CCAAACCGAATCTTCCCAATCTTATGTTTGGCTGCAGCAAGATGCTCCAC	528
Db	205	SerPheProValGluIleLeuProPheLeuTyrIleGlyCysAlaIAspSerThrAsn	224
QY	529	AAGAGCTGATACACAGCAAGATGGATTGTTATGTTGTTAAAGCCAGCTATACCTGTCCA	588
Db	225	LeuAspValLeuGluGluPheGlyTyrIleTyrIleLeuAsnAlaThrProAsnLeuPro	244
QY	589	AAGCCT--GACTTATCCCGAGTCTCAATTCCTGCTGCTGCTGATGAATGACAGCTTT	645
Db	245	AsnLeuPheGluAlaGluIleGlyGluPheTyrTyrGluGlnIleProIleSerAspHisTirp	264
QY	646	TCGTGAAGAAATTTTGCCTGGTGTGCAACAATGATAGATTCATCATGGAAGCAAAAGCC	705
Db	265	SerGlnAsnLeuSerGlnPhePheProGluAlaIleSerPheIleAspGluAlaArgGly	284
QY	706	TCGAATGATGTGTTCTAAGTGCACGTGTTAGTGGGATCTCCGCTCGCACCATGCGCT	765
Db	285	LysAsnCysGlyValLeuValHisCysLeuAlaGlyIleSerIAspSerValThrValThr	304
QY	766	ATGCCCACTCACTGAAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAA	825
Db	305	ValAlaIleTyrIleuMetIleuLeuAsnLeuSerMetAsnAspAlaTyrAspIleValLys	324
QY	826	GAAGAAAGACCTACTATATCTTCCAAACTCAATTTTCTGGGCCAATCCTGTGACATAGG	885
Db	325	MetIleTyrSerAsnIleSerProAsnPheAsnMetIleGluIleuLeuAspPheGlu	344
QY	886	AACAAATTT 894	
Db	345	ArgThrIleu 347	

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Protein-tyrosine phosphatase vhp-1 (EC 3.1.3.48).
 GN vhp-1 OR P08B1.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Briscot N2;
 RA Chisoe S.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- Non-receptor class dual specificity subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL/ U23178; AAC6719.1; -.
 DR PIR; T15969; T15969.
 DR HSSP; Q16828; IMKP.
 DR WormRep; P08B1.1; CE01899.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR SMART; SM00195; DSPC; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KM HydroLase.
 FT DOMAIN 200 269 PROTEIN-TYROSINE PHOSPHATASE.
 FT DOMAIN 92 95 POLY-SER.
 FT DOMAIN 351 354 POLY-SER.
 FT DOMAIN 465 472 POLY-SER.
 FT DOMAIN 483 488 POLY-SER.
 FT ACT SITE 224 224 PHOSPHOCYSTEINE INTERMEDIATE (BY
 FT SIMILARITY).
 SQ SEQUENCE 619 AA; 66354 MW; 369B326F615D0529 CRC64;
 Alignment Scores:
 pred. No.: 5.09e-34 Length: 619
 Score: 450.00 Matches: 90
 Percent Similarity: 69.66% Conservative: 34
 Best Local Similarity: 50.56% Mismatches: 48
 Query Match: 27.66% Indels: 6
 DB: 1 Gaps: 4
 US-10-029-345a-108_COPY_538_1443 (1-906) x VHP1_CABEL (1-619)
 QY 367 GGTGGTTGCTGAGTCTCTCGTTGTTCCCTGCGCTCTGTGAGAAATCC----- 420
 DB 99 GtGtGtPhelYsgInhEaIngaIngnInYrPrGInLeuCySeInuSerGInuGtYmEe 118
 QY 421 ACTTACTGCTTACCTGATCTTCTGACGCTTGCTTA-----CCTGTGCGCAACATTGGG 474
 DB 119 ThArJleuPrGInSerInuSerGInPrGyLeuSerGInPrGtRgIyaSp---GtY 137
 QY 475 CCAACCGAATTTCTCCATCTTATCTTGCTGCGCAGGAGGATGCTCCACAGAGAG 534
 DB 138 lIeThrLeuIeThrPrGInSerInuSerGInuGtYmEe 157
 QY 535 CTGATACAGAGAGATGGATTGTTATGTTAAATCCAGTATACCTGTCCAAAGCCT 594
 DB 158 MeTleuYaPhAlaLeuYaSpIleSerValIleAsnLeuSerMetThrCySpIleYSer 177

QY 595 GACTTATCCCGAG---TCATATTCGCGGTCGCTGTGAATGACGCTTTGAG 651
 DB 178 ValYleYsgInhEaIngaIngnInYrPrGInLeuCySeInuSerGInuGtYmEe 197
 QY 652 AAAATTTGCGCGTGGTGGCAATATGATTTGATTTGAGAAAGCAAAAGCTCCAT 711
 DB 198 lYsLeuSerPrGtYrThPrGInSerInuSerGInuGtYmEe 217
 QY 712 GGATGTTCTTACTGACATGTTTACTGCGATCTCCGCTCGCCACCATGCTTATGCC 771
 DB 218 lYsLeuCySeInuSerGInuGtYmEe 237
 QY 772 TACATATGAGAGATGATGACATGCTTATGATGAAAGCTTACGATTTGAGAAAGAAA 831
 DB 238 TyRleMeCaRgYrTyRleYsMeGtYsEaPaPaIaTyRgYrValYsGtYmEe 257
 QY 832 AGACACTATATCTCCAAACTCAATTTTCTGCGCAACTCTGACATATGAG 885
 DB 258 ArgPrGInSerPrGInSerPrGInSerPrGInSerPrGInSerPrGInSerPrGIn 275
 RESULT 10
 DUSA CHICK STANDARD; PRT; 375 AA.
 AC 09PW71;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
 DE phosphatase-2) (MKP-2).
 GN DUSP4 OR MKP2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NC NCB1_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White Leghorn;
 RX MEDLINE=20379359; PubMed=10918612;
 RA Fu S.-L., Waha A., Vogt P.K.;
 RT Identification and characterization of genes upregulated in cells
 RT transformed by v-jun.";
 RL Oncogene 19:3537-3545(2000).
 CC -1- FUNCTION: Regulates mitogenic signal transduction by
 CC dephosphorylating both Thr and Tyr residues on MAP kinases ERK1
 CC and ERK2 (By similarity).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF167296; AAD4656.1; -.
 DR HSSP; Q16828; IMKP.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.

CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U21108; AAA85119.1; -;
CC EMBL; U48807; AAC50452.1; -;
CC EMBL; BC002671; AAH02671.1; -;
CC EMBL; BC014565; AAH14565.1; -;
CC HSSP; Q16828; IMKP.
CC DR Genew; HGNC:3070; DUSP4.
CC DR MIM; 602747; -;
CC DR GO; GO:0005634; C:nucleus; TAS.
CC DR GO; GO:0008330; F:protein tyrosine/threonine phosphatase acti. . .; TAS.
CC DR GO; GO:000165; P:MAPKK cascade; TAS.
CC DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
CC DR InterPro; IPR000340; DS_phosphatase.
CC DR InterPro; IPR008343; MAPK_phosph.
CC DR InterPro; IPR001763; Rhodanese-like.
CC DR InterPro; IPR000387; TYR_phosphatase.
CC DR Pfam; PF00782; DSPC; 1.
CC DR Pfam; PF00581; Rhodanese; 1.
CC DR PRINTS; PRO1764; MAPKPHPTASE.
CC DR SMART; SM00195; DSPC; 1.
CC DR SMART; SM00450; RHOD; 1.
CC DR PROSITE; PSS0206; RHODANSE_3; 1.
CC DR PROSITE; PSS00383; TYR_PHOSPHATASE_1; 1.
CC DR PROSITE; PSS0056; TYR_PHOSPHATASE_2; 1.
CC DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
CC KW HydroLase; Nuclear protein.
CC FT DOMAIN 41 159 RHODANSE.
CC FT DOMAIN 197 394 PROTEIN-TYROSINE PHOSPHATASE.
CC FT ACT_SITE 280 280 PHOSPHOCYSTEINE INTERMEDIATE (BY
CC SIMILARITY).
CC FT CONFLICT 111 111 R -> G (IN REF. 2).
CC SQ SEQUENCE 394 AA; 42953 MW; 060397175986952E CRC64;

Alignment Scores:
Pred. No.: 6.65e-32 Length: 394
Score: 427.00 Matches: 102
Percent Similarity: 53.85% Conservative: 66
Best Local Similarity: 32.69% Mismatches: 113
Query Match: 26.24% Indels: 31
DB: 1 Gaps: 9

US-10-029-345A-108_COPY_538_1443 (1-906) x DUSP4_HUMAN (1-394)
QY 45 GGTGGCTCTGCTGGAAGTGAACGAA-----AAAGT 77
DB 26 G|G|V|A|L|A|G|I|Y|S|E|R|I|S|E|R|H|L|E|U|G|I|Y|L|E|U|P|R|O|S|E|R|I|Y|L|Y|S|C|Y 45
QY 78 GCTGCTAATGATGACCGGCGCATTTGTGAAATCAATACATCCACATTTTGAAGCCAT 137
DB 45 s|e|u|e|u|l|e|u|a|b|p|c|y|a|b|g|p|r|o|p|h|e|u|a|l|a|h|s|e|r|a|l|g|l|y|r|i|l|e|u|g|i|y|s|e|r|v|a 65
QY 138 TAAATATCAATGCTCCCAAGTTATGAAGCAAGTTTSCAAAGCAAAATGTTAATTAC 197
DB 65 l|a|n|v|a|l|a|y|c|y|a|b|a|n|h|r|i|l|e|v|a|l|a|y|a|g|a|y|g|-|A|l|d|y|s|e|r|v|a|l|e|u|g|i 84
QY 198 AGAGCTCATC-----CAGCATTCAGGAAACATTAAGGTTGACATGATTGACAGTCA 251
DB 198 AGAGCTCATC-----CAGCATTCAGGAAACATTAAGGTTGACATGATTGACAGTCA 251

DB 84 u|g|i|n|i|l|e|u|p|r|o|a|l|g|i|u|g|i|u|v|a|l|a|y|a|l|a|y|l|e|u|a|y|s|e|r|c|i|y|l|e|u|y|r|s|e|r|a|l 104
QY 252 GGTGTAAGTTTACGATCAAAAGCTCCCAAGATGTTGGCTCTCTTCCAGACGTTTCT 311
DB 104 a|v|a|i|l|e|v|a|i|y|a|b|p|c|y|a|b|g|l|a|y|s|e|r|p|r|o|a|l|g|l|a|g|i|u|s|e|r|l|e|u|a|y|g|i|u|s|e|r|t|h|r|v|a 124
QY 312 CACTGACTCTCGGGAAGTGGAGAGAGCTTC-----AACTGCTTCACTGCTTGC 365
DB 124 l|s|e|r|l|e|u|v|a|i|g|i|a|l|a|l|e|u|a|y|a|b|g|r|a|b|e|n|i|a|g|i|u|a|y|t|h|a|p|i|l|e|y|s|e|r|l|e|u|y 144
QY 366 AGGTGGATTTGCTGAAGTTCTCTGTTGTTTCCCTGCGCTCTGTGAAGAAATTCACCTT 425
DB 144 s|g|i|y|l|y|g|i|u|a|y|p|h|e|s|e|r|c|i|y|l|u|y|r|p|r|o|g|i|u|p|h|e|y|s|e|r|t|h|r|y|s|a|l|e 164
QY 426 A-----GTCCCTACCTGCATTTCTCAGCC- 450
DB 164 u|a|l|a|l|e|p|r|o|p|r|o|v|a|l|p|r|o|s|e|r|a|l|a|t|h|r|g|i|u|p|r|o|l|e|u|a|b|p|l|e|u|i|y|c|y|s|e 184
QY 451 ----TGC---TTACSTGTGSCCAACAT---GGGCGCAACCGCAATTTCCCAATCTTGA 500
DB 184 r|s|e|r|c|y|s|e|r|a|l|a|y|h|i|a|l|a|l|a|y|a|b|p|e|t|l|e|u|a|b|p|l|a|l|e|u|i|y|l|e|t|h|r|a|l 224
QY 501 TCTTGGCTGCGCAGAGATGCTCTCAACAAGAGCTGATACAGCAATGGAGTTGGTTA 560
DB 204 l|l|e|u|g|i|y|s|e|r|a|l|a|y|h|i|a|l|a|l|a|y|a|b|p|e|t|l|e|u|a|b|p|l|a|l|e|u|i|y|l|e|t|h|r|a|l 224
QY 561 TGCTTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTTATCCCGAGCTCATTTCC 620
DB 224 a|l|e|u|a|n|v|a|i|s|e|r|s|e|r|a|b|p|c|y|a|b|a|n|---N|i|s|p|h|e|g|i|u|y|h|i|s|T|r|c|i|n|t|y|r|l|y 243
QY 621 GCGTGCGCTGTAATGACAGCTTTGTGTGAAATTTTGGCGGTGTGGCAATTCAGT 680
DB 243 s|c|y|s|i|l|e|r|p|r|o|v|a|i|g|i|u|s|p|h|i|s|y|b|a|l|a|p|i|l|e|s|e|r|t|h|r|p|h|e|c|i|u|a|i|l 263
QY 681 AGATTTCATTGAGAAAGCAAAAGCTCCCAATGATGTTTATGTCACCTGTTTAACTGG 740
DB 263 e|g|i|u|y|r|i|l|e|a|p|a|l|a|l|y|a|b|p|c|y|a|b|g|i|y|a|g|v|a|i|l|e|u|v|a|h|i|c|y|s|i|n|a|g|i 283
QY 741 GATCTCCCGCTCGGCACCATCGCTATCGCTTATCATCATGAAGAGATGACATGCTTT 800
DB 283 y|i|l|e|s|e|r|a|y|s|e|r|a|l|a|t|h|r|i|l|e|c|y|e|u|a|l|a|y|l|e|u|m|e|m|e|t|l|y|s|a|y|a|l|a|y|g|l|e 303
QY 801 AGATGAGCTACAGATTGTGTGAAAGAAAGAAAGCACTACTATATTCSCAAATTTCAATT 860
DB 303 u|g|i|u|g|i|u|a|b|p|h|e|u|i|p|h|e|v|a|i|l|y|s|i|n|a|y|g|s|e|r|i|l|e|s|e|r|p|r|o|a|n|p|h|e|s|e|r|h 323
QY 861 TCTGGGCAACTCTGACTATGAGAAAGATT 894
DB 323 e|m|e|t|g|i|y|i|n|l|e|u|i|g|i|n|h|p|h|e|i|u|s|e|r|i|n|v|a|l 334

RESULT 12
DUST_MOUSE
ID DUST_MOUSE STANDARD; PRT; 367 AA.
AC P28563;
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
DE 3CH134) (Protein-tyrosine phosphatase ERP).
GN DUSP1 OR PTPN10 OR MKP1 OR 3CH134 OR PTPN16.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1] -
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=92158357; PubMed=1741163;
RA Charles C.H., Abler A.S., Lau L.F.;
RT "CDNA sequence of a growth factor-inducible immediate early gene and
RT characterization of its encoded protein";
RT Oncogene 7:187-190(1992).

[2]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RX MEDLINE=93360956; PubMed=8355678;
 RA Noguichi T., Metz R., Chen L., Mattei M.-G., Carrasco D., Bravo R.;
 RT "Structure, mapping, and expression of erp, a growth factor-inducible
 RT gene encoding a nontransmembrane protein tyrosine phosphatase, and
 RT effect of ERP on cell growth.";
 RL Mol. Cell. Biol. 13:5195-5205(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RP STRAIN=Czech II; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Klausner R.L., Reingold E.A., Grouse L.H., Derye J.G.,
 RA Altschul S.F., Zeeberg B., Wagner L., Shemen C.M., Schaefer G.D.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.J.,
 RA Raha S.S., Lequellano N.A., Peters G.J., Abremski R.D., Mullaly S.J.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.L., Skalska U., Smallus D.E.,
 RA Schenck J.A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=94037096; PubMed=822188;
 RA Sun H., Charles C.H., Lau L.F., Tonks N.K.;
 RT "Wk-1 (3CH134), an immediate early gene product, is a dual
 RT specificity phosphatase that dephosphorylates MAP kinase in vivo.";
 RL Cell 75:487-493(1993).
 CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
 CC kinase ERK2 on both Thr-183 and Tyr-185.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- INDUCTION: By growth factors.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X61940; CA443944.1; -;
 DR EMBL; X64851; AA827882.1; -;
 DR EMBL; BC006967; AA06967.1; -;
 DR PIR; A54681; S24411.
 DR HSP; Q16828; IMKP.
 DR MGD; MGI:105120; Dusp1.
 DR InterPro: IPR000340; DS phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR Pfam; PF00782; DSCP; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PRO1764; NAKPHPTASE.
 DR SMART; SM00195; DSCP; 1.

DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PSS0206; RHODANES_3; 1.
 DR PROSITE; PSS0383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PSS0056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolyase; Cell cycle.
 FT DOMAIN 20 137
 FT DOMAIN 175 367
 FT ACT SITE 258 258
 FT MOTIF 258 258
 SQ SEQUENCE 367 AA, 39369 MW, 5085FP0FEED19AB CRC64;
 Alignment Scores:
 Pred. No.: 1.01e-31 Length: 367
 Score: 425.00
 Percent Similarity: 50.17% Conserved: 104
 Best Local Similarity: 34.32% Mismatches: 48
 Query Match: 26.12% Indels: 129
 DB: 1 Gaps: 5
 US-10-029-345a-108_COPY_538_1443 (1-906) x DUS1_MOUSE (1-367)
 QY 43 TTGGTGGCTCTGCTGGAAGTGAACGGAAGAAAGTCTGTAATTAATGATGCGGCATT 102
 DB 13 LeuAlaArgLeuLeuArgGluGluAlaValAlaGlnCysLeuLeuLeuAspCysArgSerPhe 32
 QY 103 GTGGAATACATACATCCCATTTTGGAGCCATTAAATCACTGCTCCAGCTTTATG 162
 DB 33 PheAlaPheAspAlaGlyHisIleAlaGlySerValAsnValArgPheSerThrIleVal 52
 QY 163 AAGCGAAGTTCGACACAGACAAAGTGTAAATTCACAGCTCATCCAGCTTACGCGA 222
 DB 53 ArgArgArgAlaValysGlyAlaMetGlyLeuGlnHisIleValProAsnAlaGluLeuArg 72
 QY 223 CATTAAGTTGACATTATGACAGTCAGACAGAGCTTGATTTACAGCAAGTCCCAAGT 282
 DB 73 GlyArgLeuLeuAlaGlyAlaTyrHisAlaValAlaValLeuLeuAspGluArgSerIleSer 92
 QY 283 GTTGCTCTCTCTTTCAGACTGTTTTCACAGTCTGCTTCTGCGTAACTG-----GAG 336
 DB 93 LeuAspGlyAlaValysArgSerGlyThrLeuAlaLeuAlaAlaGlyAlaLeuCysArgGlu 112
 QY 337 AAGAGCTTCACTCTGTTACCTGCTGAGAGTGGCTTCTGAGTCTCTCTGTTGTTTC 396
 DB 113 AlaArgSerThrGlnValPhePheLeuGlnGlyGlyTyrGluAlaPheSerAlaSerCys 122
 QY 397 CCGGCGCTCTGTTGAGAAATTCACCTGATCCCTACCTGATTCATTTCTCAGGCT----- 450
 DB 133 ProGluLeuCysSerIleSerIleThr-----ProThrGlyLeuSerIleProLeuSer 150
 QY 451 -----TGCCTTACCTGTTGCCAATTT----- 471
 DB 151 ThrSerValProAspSerAlaGluSerClyCysSerSerCysSerThrProLeuTyrAsp 170
 QY 472 -----GGGCGCAACCGCAATTTCTCCAAATCTTATCTTGCTGCGCAGAGATGTCCTC 525
 DB 171 GlnGlyGlyProValGluIleLeuSerPheLeuTyrLeuGlySerIleAlaTyrHisAlaSer 190
 QY 526 AACCAAGAGCTGATACAGCAGAAATGGATGTTATGTTGTTAAATGCCAGCTTACCTGT 585
 DB 191 ArgIysAspMetLeuAspAlaLeuGlyIleThrAlaLeuIleAsnValSerIleAsnCys 210
 QY 586 CCAAAAGCTGACTTATCCCGAGTCACTTCTCGCGTGCCTGCTGTAATGACAGCTTT 645
 DB 211 ProAsn---HisPheGlnGlyHisTyrGlnTyrIleSerIleProValGluAspHis 229
 QY 646 TGTGAGAAATTTTGGCGGTGGTGGACAAATCATGATGATTTCAAGAAAGCAAAAGCC 705
 DB 230 LysAlaAspIleSerSerTrpPheAsnGluAlaIleAspPheIleAspSerIleTyrAsp 249
 QY 706 TCCAAATGATGTTGTTAGTCACTTTAGCTGGAGATTCCTCGCTCCGACACATGCTGCT 765
 DB 250 AlaGlyGlyArgValPheValHisCysGlnAlaGlyIleSerArgSerAlaThrIleCys 269

```

QY 766 ATGCCCATCATGAGAGATGACATGCTTTAGATGAGCTTACAGATTTTGAAA 825
Db 270 LeuAlaIyrLeuNeuAcGlnThrAsnArgValIyrLeuAspGluAlaPheGluPheValIys 289
QY 826 GAAAAAGACCTATATATCTCCAAACTTATTTTGAGCCCAACTCTGATGATAG 885
Db 290 GlnArgGserLeuIleIleSerProAsnPheSerPheMetGlyGlnLeuLeuGlnPheGlu 309
QY 886 AAGACATT 894
Db 310 SerGlnVal 312

RESULT 13
DUS4_RAT
ID DUS4_RAT STANDARD; PRT; 395 AA.
AC 062767;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
DE phosphatase-2) (MKP-2).
DE DUSP4 OR MKP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Phenochromocytoma;
RX MEDLINE=95301550; PubMed=7782322;
RA Miera-Press A., Rlm C.S., Yao H., Robertson M.S., Stork P.J.S.;
RT "A novel mitogen-activated protein kinase phosphatase. Structure,
RT expression, and regulation."
RL J. Biol. Chem. 270:14587-14596(1995).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases ERK1
CC and ERK2 (by similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed at moderate levels in nearly all
CC tissues and cells including brain, spleen, and testes with the
CC higher expression in the heart and lung and lower expression in
CC skeletal muscle and kidney. Undetectable in liver. Expressed in
CC many areas of the brain with very strong expression in the
CC hippocampus, piriform cortex, and the suprachiasmatic nucleus.
CC -1- INDUCTION: By mitogens and by stress.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U23438; AAC52493.1; -.
CC HSSP: Q16828; IMKP.
CC InterPro: IPR000340; DS_phosphatase.
CC InterPro: IPR008343; MAPK_phosph.
CC InterPro: IPR001763; Rhodanese-like.
CC InterPro: IPR000387; TYR_phosphatase.
CC Pfam: PF00782; DSpPc; 1.
CC Pfam: PF00581; Rhodanese; 1.
CC PRINTS: PR01764; MAPKPHPTASE.
CC SMART: SM00195; DSpC; 1.

```

```

DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE 3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE 1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE 2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW HydroLase; Nuclear protein.
FT DOMAIN 42 160 RHODANSE.
FT DOMAIN 198 395 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 281 281 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SEQUENCE 395 AA; 43187 MW; A90BFFD378A050FD CRC64;
SIMILARITY).

Alignment Scores:
Pred. No.: 1 41e-31 Length: 395
Score: 423.50 Matches: 98
Percent Similarity: 52.30% Conservative: 61
Best Local Similarity: 32.24% Mismatches: 118
Query Match: 26.03% Indels: 27
DB: Gaps: 7

US-10-029-345a-108_COPY_538_1443 (1-906) x DUS4_RAT (1-395)
QY 46 GTGGCTGTGTGGAAAGTGAAGCGAAAGTGTCTATTTAGATGAGCCGCAATTTGTG 105
Db 38 LeuGlyLeuLeuSerGly-----LysCysLeuLeuLeuSerCysArgProPheLeu 55
QY 106 GAATACAAATACATCCACATTTTGGAAAGCAATTAATACATGCTCCAAAGTTATGAG 165
Db 56 AlaHisSerAlaGlyTyrIleArgGlySerValAsnValArgCysAsnThrIleValArg 75
QY 166 CGAAGGTGAACAGAGCAAGCAAGTTAATACAGAGCTAC-----CAGCATTCAGCG 219
Db 76 ArgArg---AlaValSerValSerLeuGlnGlnIleuProAlaGluGluVal 94
QY 220 AACATAGATGATGACATGATGATGACATGATGATGATGATGATGATGATGATGATGATG 279
Db 95 ArgAlaArgLeuArgSerGlyLeuTyrSerAlaValIleValIleArgIleArgSerPro 114
QY 280 GATGTGCTGTCTCTCTTCAAGATGTTTTCATGATGATGATGATGATGATGATGATGATG 339
Db 115 ArgAlaGluSerLeuArgGluLysSerThrValSerLeuValAlaGlnAlaValArg 134
QY 340 AGCTTC-----AAGCTGTGACCTGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 393
Db 135 AsnAlaGlnArgThrAspIleCysValLeuLysGlyGlyTyrGluArgPheSerSerGlu 154
QY 394 TTCCCTGAGCTGTGTGAAGAAATCCAGTACGTCCTTACGTCGATTTCTGACCTTGC 453
Db 155 TyrProGluPheCysSerIleThrIleValAlaLeu-----AlaAlaIleProProVal 172
QY 454 TTACCTGTGTCGAACATT----- 471
Db 173 ProProSerThrAsnGluSerLeuSerLeuGlyCysSerSerCysGlyThrProLeuIle 192
QY 472 -----GGGCGAACCGAATTTCTCCCAATCTTTATCTTCTGAGCGCGAAGATGTC 522
Db 193 AspGlnGlyIleProAlaIleuProPheLeuTyrLeuGlySerAlaTyrGlnVal 212
QY 523 CTCACAAAGACATGATGACAGAGATGATGATGATGATGATGATGATGATGATGATGATG 582
Db 213 AlaArgArgAspMetLeuAspAlaLeuGlyIleThrAlaLeuLeuAsnValSerSerAsp 232
QY 583 TGTCCAAAGCTGACTTATCCCGAGTCAATTTCTGCGGTGCGGTGCGGTGATGACAG 642
Db 233 CysProAsn---HisPheGluGlyIleTyrGlnTyrGlyCysValIleProValGluAsp 251
QY 643 TTTTGGAAATAATTTGCGCGGTGAGCAATGATGATGATGATGATGATGATGATGATGATG 702
Db 252 HisValAlaAspIleSerSerThrPheMetGluAlaIleGlnTyrIleAspValIys 271
QY 703 GCGTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 762
Db 272 AspCysArgGlyArgValLeuValHisCysValAlaGlyIleSerArgSerIleThrIle 291

```



```

QY 763 GCTATGCTTACATCATGAGAGATGCATGCTTATGATGAGCTTACAGATTTCG 822
Db 292 CysleuAlaIyrleuMetMetIysIysArgValAlaArgleuGluGluAlaPheGluPheVal 311
QY 823 AAGAGAAAAAGACCTACTATATCTCCAACTTCAATTTTCGGGCCCACTCTGAGCTAT 882
Db 312 LysGlnArgArgSerIleIleSerProAsnHeserPheMetGlyGlnleuGlnPhe 331
QY 883 GAGAGAAGAT 894
Db 332 GluSerGlnVal 335

RESULT 14
DUS2_HUMAN
ID DUS2_HUMAN STANDARD; PRT; 314 AA.
AC 005923;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 2 (EC 3.1.3.48) (EC 3.1.3.16)
DE DUSP2 OR PAC1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=93206122; PubMed=7681221;
RA Rohan P., Davis P., Moskaluk C.A., Kearns M., Kruttsch H.,
RA Stebenlist U., Kelly K.,
RT "PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.",
RL Science 259:1763-1766 (1993).
RN 12;
RP SEQUENCE FROM N.A.
RX MEDLINE=96070437; PubMed=7590752;
RA Yi H., Morton C.C., Meremowicz S., McBride O.W., Kelly K.;
RT "Genomic organization and chromosomal localization of the DUSP2 gene,
RT encoding a MAP kinase phosphatase, to human zp11.2-q11.",
RL Genomics 28:92-96 (1995).
RN 3;
RP SEQUENCE FROM N.A.
RX TISSUE=B-cell;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein W., Soares M.B., Bonaldi M.F., Casavant T.D., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Toshiyuki S., Carninci P., Prange C.,
RA Bobak S.S., McMan P.J., McKernan K.J., Adams R.D., Mulaly S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Rodriguez S.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Schermer A., Schein J.E., Jones S.J.W., Skalek U., Smolius D.E.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases
CC ERK1 and ERK2.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear.

```

```

CC -1- TISSUE SPECIFICITY: In hematopoietic tissues.
CC -1- INDUCTION: By mitogens.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, J11329; AAA50779.1; -.
DR EMBL, U23853; AAA86112.1; -.
DR EMBL, BC007771; AAH07771.1; -.
DR PIR, A57126; A57126.
DR PDB, 1IKZ; 3D-MAY-02.
DR Genew; HGNC:3068; DUSP2.
DR MIM; 603068.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; TAS.
DR GO; GO:0008330; F:protein tyrosine/threonine phosphatase acti. . .; TAS.
DR GO; GO:0000188; P:inactivation of MAPK; TAS.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK-phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; RHOD; 1.
DR SMART; SM00450; DSPC; 1.
DR PROSITE; PS50206; RHODANES 3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase; Nuclear protein; 3D-structure.
FT DOMAIN 23 144 RHODANES.
FT FT 237 302 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 257 257 PHOSPHOCYSTEINE INTERMEDIATE (BY
SQ SEQUENCE 314 AA; 34399 MW; FDD3543C6DE10C45 CRC64;
SIMILARITY)
Alignment Scores:
Pred. No.: 4,43e-31
Score: 418.00 Length: 314
Percent Similarity: 50.83% Matches: 109
Best Local Similarity: 35.97% Conservative: 45
Query Match: 25.69% Mismatches: 105
Db: 1 Indels: 44
Gaps: 8
US-10-029-345a-108_COPY_538_1443 (1-906) x DUS2_HUMAN (1-314)
QY 70 GAAAAAGTCTGCTTATGATGATGAGCGGCATTTGTGATATACATATCCAGATTTCG 129
Db 25 GluArgThrIleuLeuAspCysArgProPheIleuAlaPheCysArgArgIleValArg 44
QY 130 GAAGCATTAATTAATCACTGCTCCAGCTTATGACGAGAGG----- 171
Db 45 AlaAlaArgProValProTTPPhaAlaIleuIleuArgArgAlaArgGlyProProAla 64
QY 172 -----TTGCAACAGACGAAAGTGTATTATGAGAGCTATCCAGATTC 216
Db 65 AlaValIleuAlaCysIleuLeuProAspArgAlaIleuArgThrArgIleuValArgGlyGlu 84
QY 217 GCGAACAATAGGTGACATTGTCACAGAAAGTTGATTGACCAAAAGCTCC 276
Db 85 Leu-----AlaArgAlaValIleuAspGluGlySer 95
QY 277 CAAGATGTCCTCTCTCTTTCAGAC-----TCTTTCTCATGTAAGTTCG 324

```

[illegible]

```

RX MEDLINE:22388257; PubMed:12477932;
RA Straubeberg R.L., Feinberg E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
RA Brownstein M.J., Udell T.B., Tothiyaki S., Carninci P., Prange C.,
RA Rana S.S., Lequand N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gudarath P.H.,
RA Richards S., Wozley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whaley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttefield Y.S.N., Krzywinski M.I., Skalak U., Smillius D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
CC kinase ERK2 on both Thr-183 and Tyr-185.
CC -1- CATALYTIC ACTIVITY: protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- INDUCTION: By oxidative stress and heat shock.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68277; CAA8438.1; -.
DR EMBL: BC022463; AAH22463.1; -.
DR EMBL: F18; S29090; S29090.
DR HSP: Q16828; IMKP.
DR Genew: HGNC:3064; DUSP1.
DR MIM: 600714; -.
DR GO: GO:0004726; F:non-membrane spanning protein tyrosine phos. . .; TAS.
DR GO: GO:0006979; P:response to oxidative stress; TAS.
DR InterPro: IPR000340; DS_PhoPhatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS: PRO1764; MAPKPHPTASE.
DR SMART: SMO0195; DSPC; 1.
DR SMART: SMO0450; RHOD; 1.
DR PROSITE: PSS50266; RHODANSE_3; 1.
DR PROSITE: PSS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PSS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PSS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydroxylase; Cell cycle.
FT DOMAIN 20 137 RHODANSE.
FT FT 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 367 AA; 39297 MW; 11BD1D39A9FCD51F CRC64;

Alignment Scores:
Pred. No.: 8,69e-31 Length: 367
Score: 415.00 Matches: 102
Percent Similarity: 49.50% Conservative: 47
Percent Local Similarity: 33.89% Mismatches: 134

```

US-10-029-345A-108_COPY_538_1443 (1-906) x DUS1_HUMAN (1-367)

[illegible]

Search completed: June 21, 2004, 12:35:00
Job time : 17.5253 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:25 ; Search time 46.8459 Seconds
(without alignments)
12204.244 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 1627

Sequence: 1 argsgccatgagatgatg.....agaagataagaaccagact 906

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xip
-Q/cgn2 1/USPTO.spool.p/US10029345/runat.21062004.122816.4193/app.query.fasta.1.3278
-DB=SPRMBL.25 -OFMT=fasta -SUFFIX=rspt -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345_QCEN.1.168_@runat.21062004.122816.4193 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPRMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	DB ID	Description
1	1552	95.4	662	Q81VT8	Q81VT8 homo sapien

2	1545	95.0	665	4	Q96N49	Q96N49 homo sapien
3	1517	93.2	660	11	Q92OR2	Q92OR2 mus musculus
4	1514	93.1	677	11	Q99MG6	Q99MG6 mus musculus
5	1377.5	84.7	355	11	Q9AG16	Q9AG16 mus musculus
6	1312	80.6	616	4	Q8NST1	Q8NST1 homo sapien
7	1172.5	72.1	622	11	Q99MG5	Q99MG5 mus musculus
8	1036	63.7	300	11	Q9AG15	Q9AG15 mus musculus
9	989	60.8	625	4	Q86SS8	Q86SS8 homo sapien
10	985	60.5	665	11	Q7RSZ9	Q7RSZ9 mus musculus
11	891	54.8	136	11	Q85ZE4	Q85ZE4 mus musculus
12	653	40.1	206	4	Q96QS2	Q96QS2 homo sapien
13	563.5	34.6	657	5	Q8ST19	Q8ST19 caenorhabditis
14	472.5	29.0	367	13	Q7SZF3	Q7SZF3 brachydanio
15	469	28.8	483	11	Q8R3L3	Q8R3L3 mus musculus
16	465	28.6	382	13	Q7R2L9	Q7R2L9 gallus galli
17	450	27.7	606	5	Q8ST18	Q8ST18 caenorhabditis
18	447	27.5	383	13	Q8BW48	Q8BW48 fugu rubripes
19	440.5	27.1	368	4	Q8NFJ0	Q8NFJ0 homo sapien
20	436	26.8	382	13	Q7R2L8	Q7R2L8 brachydanio
21	430	26.4	362	13	Q803B2	Q803B2 brachydanio
22	427	26.2	411	4	Q13549	Q13549 homo sapien
23	423.5	26.0	378	13	Q91663	Q91663 xenopus laevis
24	423.5	26.0	398	11	Q8BFV3	Q8BFV3 mus musculus
25	421.5	25.9	369	13	Q91790	Q91790 xenopus laevis
26	419.5	25.8	369	13	Q90W58	Q90W58 xenopus laevis
27	416	25.6	367	11	Q64193	Q64193 rattus norvegicus
28	413	25.4	367	11	Q80ZM1	Q80ZM1 mus musculus
29	413	25.4	367	11	Q63683	Q63683 rattus norvegicus
30	411	25.3	439	5	Q81G35	Q81G35 caenorhabditis
31	403.5	24.8	353	13	Q42253	Q42253 gallus gallus
32	403.5	24.8	360	13	Q7ZVL8	Q7ZVL8 brachydanio
33	403	24.8	436	11	Q99KC2	Q99KC2 mus musculus
34	397	24.4	452	11	Q8K1S9	Q8K1S9 mus musculus
35	397	24.4	452	11	Q7TNL7	Q7TNL7 mus musculus
36	382.5	23.5	354	4	Q8N4A4	Q8N4A4 homo sapien
37	347.5	21.4	411	5	Q9VSV1	Q9VSV1 drosophila
38	347.5	21.4	411	5	Q9VSV5	Q9VSV5 drosophila
39	325	20.0	303	4	Q9NSM1	Q9NSM1 homo sapien
40	306	18.8	177	11	Q9CSL5	Q9CSL5 mus musculus
41	287	17.6	369	5	Q44128	Q44128 caenorhabditis
42	285.5	17.5	476	5	Q46122	Q46122 drosophila
43	280.5	17.2	46	5	Q9VHV8	Q9VHV8 drosophila
44	272	16.7	290	5	Q86P14	Q86P14 drosophila
45	269	16.5	241	5	Q9VWV4	Q9VWV4 drosophila

ALIGNMENTS

RESULT 1
Q81VT8 PRELIMINARY; PRT; 662 AA.
ID Q81VT8
AC Q81VT8
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to dual specificity phosphatase 16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAUSBERG R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC042101; AA042101.1; -
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPc; 1.


```

QY 121 CACATTTTGGAGCCATTATATCAATGCTCCAGCTTATGAAAGGAGTTGCAACAG 180
Db 41 Hieileuenglunlaileanilleanencyserilyseuuectlybargagluengln 60
QY 181 GACAAAGTGAATTAATACAGAGCTCATCCAGCATTCAGCAAGAAATAAGTTGACATTGAT 240
Db 61 AspLyvalleuilethrgluileuileglnhiserilaishlybvalaplleasp 80
QY 241 TGCAGTCAGAGGTTGATTGATTAAGATCAAAAGCTCCCAAGATGGCTCTCTCTTCA 300
Db 81 CyssercInlyvalValValValYrAspGlnserSerGlnAspValAlaSerleuSerSer 100
QY 301 GACTGTTTTCTCCTGACTGTTCTGCGGTAATCGGAGAGAGCTTCAACTGTCTACCTG 360
Db 101 AspCyphelueuthrValleuileuglyLyseuglnlysserPheanserValHisleu 120
QY 361 CTTCGAGAGTGGTTGCTGAGTTCTCTCGTTGTTCCCTGGCCCTCTGTGAAGAAATCC 420
Db 121 leuAlaglylyPheAlaglnPheSerArgCypheroillyleuCyseuglylysser 140
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCTTGTCTTACCTGTGCAACATTTGGCCCAAC 480
Db 141 ThrleuValProthrCyseileSerGlnProCyseleuProValAlaAenilleglyProthr 160
QY 481 CGAATCTTCCCAATCTTTATCTTGGCTGCGCAGCAGAGATGCTCTCAACAGAGCTGATA 540
Db 161 leuileuProanleuuryleuglyCyseGlnArgpValleuanlysgluLeuMet 180
QY 541 CAGCAGAAATGGATGGTATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 600
Db 181 GlnGlnanelllylleglylyrvalleuanhAlaseranthrcySprolySprohepPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGCGTGTGCTGTGAATGACAGCTTTGTGAAAAATTTTG 660
Db 201 IleProGlnserHisPheleuArgValProValAsnSerPheCyseGlnlySleu 220
QY 661 CCGTGTGTGCAAAATGATGATTTGATTTGAGAAAGCAAAAGCTTCATGATGTT 720
Db 221 ProtrleuAspLySerValAspPheIleGlnlySAlaAspSerAnGlyCyseVal 240
QY 721 CTAGTCACTGTTAGTGGGATCTCCCGCTCCGCCACATCGCTATCGCTACATCATG 780
Db 241 leuValHisCyseleuAlaglyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIle 260
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT 840
Db 261 LyArgMetAspMetSerleuAspGlnAlaIlyrArgPheVallysglnlySArgProthr 280
QY 841 ATATCTTCAAACTTCAATTTTTCGGGCAACTCTGAGACTATGAGAAAGATTAGAAC 900
Db 281 IleSerProanPheanPheleuGlyGlnleuAspPyrGlnlySlySlelySAsn 300
QY 901 CAGACT 906
Db 301 GlnThr 302

```

RESULT 3

```

Q920R2 PRELIMINARY; PRT; 660 AA.
ID Q920R2;
AC Q920R2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MAP kinase phosphatase-7.
GN DUSP16 OR 3830417M17RIK OR MKP-7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;

```

```

RX MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
RT "MKP-7, a Novel Mitogen-activated Protein Kinase Phosphatase,
RT Functions as a Shuttle Protein.";
RL J. Biol. Chem. 276:39002-39011 (2001).
DR EMBL; AB052157; BAB47240.1; -.
DR HSSP; Q16828; MKP.
DR MGD; MGI:1917936; Dusp16.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0005634; Cytosol; IDA.
DR GO; GO:0005515; Fibrinogen binding; IPI.
DR GO; GO:0000188; P-inactivation of MAPK; IDA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; Tyr_kinase.
DR Pfam; PF00782; DsPc; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DsPc; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS02026; RHODANASE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SQ SEQUENCE 660 AA; 72695 MW; DB609FCDADAA309 CRC64;

Alignment Scores:
Pred. No.: 3,6e-147 Length: 660
Score: 1517.00 Matches: 290
Percent Similarity: 98.34% Conservative: 7
Best Local Similarity: 96.03% Mismatches: 5
Query Match: 93.24% Indels: 0
Gaps: 0

US-10-029-345A-108_COPY_538_1443 (1-906) x Q920R2 (1-660)
QY 1 ATGCCCATGACGATGATTTGGAATCTCAATTTGTTACTGAGAGGTTGGTGCTTCTCGAA 60
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnSerleuValAlaLeuGln 20
QY 61 AGTGAAGGAAAGAGCTGCTAATTTGATGAGCGGCACTTTGTGAAATCAATATCATCC 120
Db 21 SerGlyThrGlnlySAlaLeuileuileaspSerArgProheValGlnlyrAnThrSer 40
QY 121 CACATTTTGGAGCCATTATATCAATGCTCCAGCTTATGAAAGGAGTTGCAACAG 180
Db 41 Hieileuenglunlaileanilleanencyserilyseuuectlybargagluengln 60
QY 181 GACAAAGTGAATTAATACAGAGCTCATCCAGCATTCAGCAAGAAATAAGTTGACATTGAT 240
Db 61 AspLyvalleuilethrgluileuileglnhiserilaishlybvalaplleasp 80
QY 241 TGCAGTCAGAGGTTGATTGATTAAGATCAAAAGCTCCCAAGATGGCTCTCTCTTCA 300
Db 81 CyssercInlyvalValValValYrAspGlnserSerGlnAspValAlaSerleuSerSer 100
QY 301 GACTGTTTTCTCCTGACTGTTCTGCGGTAATCGGAGAGAGCTTCAACTGTCTACCTG 360
Db 101 AspCyphelueuthrValleuileuglyLyseuglnlysserPheanserValHisleu 120
QY 361 CTTCGAGAGTGGTTGCTGAGTTCTCTCGTTGTTCCCTGGCCCTCTGTGAAGAAATCC 420
Db 121 leuAlaglylyPheAlaglnPheSerArgCypheroillyleuCyseuglylysser 140
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCTTGTCTTACCTGTGCAACATTTGGCCCAAC 480
Db 141 ThrleuValProthrCyseileSerGlnProCyseleuProValAlaAenilleglyProthr 160
QY 481 CGAATCTTCCCAATCTTTATCTTGGCTGCGCAGCAGAGATGCTCTCAACAGAGCTGATA 540
Db 161 ArgIleuProanleuuryleuglyCyseGlnArgpValleuanlysgluLeuMet 180

```



```

OY 541 CAGGAGATGAGATTGGTTATGTGTTAAATGCCACTATACCTGTGCAAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerAnthrCysProIysProAspPhe 200
OY 601 ATCCCGAGCTCTCATTTCTGCGTGCCTGCGTGAATGACACTTTTGTGGAATAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIysIleLeu 220
OY 661 CCGTGGTTGGAACAATTCAGTATGATTCATTGAGAAAGCAAAAGCTTCAATGATGTGTT 720
DB 221 ProTProLeuAspIysSerValAspPheIleGluIysAlaIysAlaSerAnthrCysVal 240
OY 721 CTAGAGCACTGTTTGGCTGGAGATCTCCGCTCCGCGCCACTGCTGTTCCCTTACATG 780
DB 241 LeuIleHisCysLeuAlaGlyIleSerArgSerHisThrIleAlaIleAlaIleAlaIleMet 260
OY 781 AAGAGATGACATGCTTTTGAATGAGCTTACAGATTGTGAAAGAAAAAGAAAGCTTACT 840
DB 261 LysArgMetAspMetSerIleAspGlnAlaIyArgPheValIysGluIysArgProThr 280
OY 841 ATATCTCCAAACTTCATTTTCTGGGCGCAACTCTGAGCTATGAGAGAAGATTAAAGAC 900
DB 281 IleSerProAsnPheAsnPheMetGlyGlnLeuMetAspIyGluIysThrIleAsnAsn 300
OY 901 CAGACT 906
DB 301 GlnThr 302

```

RESULT 4

```

OY 99MG6 PRELIMINARY; PRT; 677 AA.
AC 099MG6;
DT 01-JUN-2001 (TRENBLREL. 17, Created)
DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
DE Map Kinase phosphatase-W A1 isoform.
GN DUSP16 OR 3830417M17R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Matsuyuchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RT Yoshikai Y.;
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages."
RL Mol. Cell. Biol. 20:6999-7009(2001).
DR EMBL: AF345951; AAK35052.1; -.
DR HSSP: Q16828; IMKP.
DR MGD: MGI:1917936; Dusp16.
DR GO: GO:0005737; Cytoplasm; IDA.
DR GO: GO:0005634; Cytoplasm; IDA.
DR GO: GO:0005515; F-protein binding; IPI.
DR GO: GO:0001088; P-kinetase activation of MAPK; IDA.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR001763; MAPK_phosph.
DR InterPro: IPR000387; Tyr_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS: PR01764; MAPKPHNTASE.
DR SMART: SM00450; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE 1; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE 2; 1.
DR PROSITE: PS00054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SQ SEQUENCE 677 AA; 74550 MW; 8B6D5B7096CEC2FC CRC64;

```

Alignment Scores:

Pred. No.:	7.37e-147	Length:	677
Score:	1514.00	Matches:	289
Percent Similarity:	98.34%	Conservative:	8
Best Local Similarity:	95.70%	Mismatches:	5
Query Match:	93.05%	Indels:	0
DB:	11	Gaps:	0

US-10-029-345a-108_COPY_538_1443 (1-906) x 099MG6 (1-677)

```

OY 1 ATGCGCCATGAGATGATTTGAAATCTCAATTGTTACTGAGAGTGTGGCTGCTTGGAA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnSerLeuValAlaLeuGlu 20
OY 61 AGTGAACGAAAAAGTCTGCTTAATTGATGACCGGCACTTTGTGGAATAATCATC 120
DB 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluIyranThrSer 40
OY 121 CACATTTTGAAGCCATTAAATATCACTGCTCCAAAGCTTATGAAAGGAAAGTTCACAG 180
DB 41 HisIleGluGlnAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
OY 181 GACAAAGTGTAAATTCAGAGCTCATTCAGCAATTCAGCAAAACATTAAGTTGACATTGAT 240
DB 61 AsparGValLeuIleThrGluLeuIleGlnHisSerAlaIySHIstIyValAspIleAsp 80
OY 241 TGCAGTCAAGAGTGTAGTTTATCATCAATCAAGCTCCAAAGATGTGCTCTGCTTCA 300
DB 81 CysAsnGlnIyValValValIyAspGlnSerSerGlnAspValGlySerIleSerSer 100
OY 301 GACTGTTTCTCATCTGATCTTCTGGGTAACGTGGAAGAAAGCTTCACTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuGluGlyIySerGlnIySerPheAsnSerValHisIleu 120
OY 361 CTTCGAGTGGGTGTGCTGCACTTCTCTGTTGTTTCCCTGCTGCTGTAAGAAATTC 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyIySerGluIySer 140
OY 421 ACTTACGCTCCATCTGATTTCTCAAGCTTGTGCTTCACTGTTGCAACATTGGGCAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
OY 481 CGAATTTCTCCAACTTTTATCTTCTGCTGCTGCAAGAGATGTCTCAAGAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuIyIyLeuGlyCysGlnArgAspValLeuValAsnIyAspLeuMet 180
OY 541 CAGGAGATGAGATTGGTTATGTGTTAAATGCCACTATACCTGTGCAAAAGCTTACT 600
DB 181 GlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerAnthrCysProIysProAspPhe 200
OY 601 ATCCCGAGCTCTCATTTCTGCGTGCCTGCGTGAATGACACTTTTGTGGAATAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIysIleLeu 220
OY 661 CCGTGGTTGGAACAATTCAGTATGATTCATTGAGAAAGCAAAAGCTTCAATGATGTGTT 720
DB 221 ProTProLeuAspIysSerValAspPheIleGluIysAlaIysAlaSerAnthrCysVal 240
OY 721 CTAGAGCACTGTTTGGCTGGAGATCTCCGCTCCGCGCCACTGCTGTTCCCTTACATG 780
DB 241 LeuIleHisCysLeuAlaGlyIleSerArgSerHisThrIleAlaIleAlaIleAlaIleMet 260
OY 781 AAGAGATGACATGCTTTTGAATGAGCTTACAGATTGTGAAAGAAAAAGAAAGCTTACT 840
DB 261 LysArgMetAspMetSerIleAspGlnAlaIyArgPheValIysGluIysArgProThr 280
OY 841 ATATCTCCAAACTTCATTTTCTGGGCGCAACTCTGAGCTATGAGAGAAGATTAAAGAC 900
DB 281 IleSerProAsnPheAsnPheMetGlyGlnLeuMetAspIyGluIysThrIleAsnAsn 300
OY 901 CAGACT 906
DB 301 GlnThr 302

```

```
RESULT 5
O9AG16 PRELIMINARY; PRT; 355 AA.
ID O9AG16;
AC O9AG16;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Map kinase phosphatase-M BI isoform.
OS DUSP16 OR 3830417M1RIK.
SN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RC Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RA Yoshikai Y.;
RT "A Novel MKP Is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages.";
RL Mol. Cell. Biol. 20:6999-7009(2001).
DR EMBL; AF345953; AAK35054.1; -.
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0001088; F:inactivation of MAPK; IDA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR00387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydrolyase; Kinase.
SQ SEQUENCE 355 AA; 39502 MW; D52C29AE21SCA285 CRC64;

Alignment Scores:
Pred. No.: 7.44e-133 Length: 355
Score: 1377.50 Matches: 272
Percent Similarity: 89.78% Conservative: 9
Best Local Similarity: 86.90% Mismatches: 20
Query Match: 84.67% Indels: 12
Gaps: 1

US-10-029-345a-108_COPY_538_1443 (1-906) x O9AG16 (1-355)
QY 1 ATGGCCCATGAGTATGATGCAATTTGTTACTAGAGGTTGGTGGCTTCTGCGAA 60
DB 1 MeclahisigluwetiegllythrglnilevalthrglnleuvalalaleuGlu 20
QY 61 AGTGAACGAGAAAAGTGTCTATATGATAGCGGCGCATTTGTGTAATCAATATCC 120
DB 21 SerGlythrglnlysvallleuileuileabserArgprohevalglutryanthSer 40
QY 121 CACATTTGGAAGCATTAATATCACTGCTCCAAAGTTATGAAGCAGAGTTGCAAG 180
DB 41 HlrlleuGlnlunlatleamilleanscyserylseuWetlysaTgaTleuGln 60
QY 181 GACAAATGTTAATTAAGCTCATTCACATTCACAGAAACTAAGTTGACATGAT 240
DB 61 AspArgValleuiletrrglnleuileglnhisserAlayshlysvaAlapleasp 80
QY 241 TGACATGAGAGGTTAGTTTACGATCAAGCCCAAGATGTTGCTCTCTTCA 300
|||||
```

```
DB 81 CysAnGlnlysvaValValTyrAspGlnSerSerGlnAspValGlySerLeuSerSer 100
QY 301 GACTGTTTCTCCTGTAATCTTGGGTTAACTGAGAAAGCTTCAACTCTGTTCACTG 360
DB 101 AspCyspheLeuThrValleuGlnGlyLyLeuGlnArgSerpheAnsSerValHsLeu 120
QY 361 CTTCGAGGTGGGTTTCTGATGCTCTCGTTGTTCCCTGAGCTGTTGTAAGAAATCC 420
DB 121 LeuAliglyValpheAlaglnpheserArgCyspheProGlyneucysgluGlySer 140
QY 421 ACTCTAGTCCCTACCTGATTTCTGACGCTTGTCTTACTGTTGCCAACATGGGCCAAC 480
DB 141 ThrleuValProThrCysIleSerGlnProCysLeuProValAlaenllleglyProThr 160
QY 481 CGAATTTCTCCCATCTTTATCTTGGCTGCGACCGAGATGTCCTCAACAAAGAGCTGATA 540
DB 161 ArgIleLeuProAsnleuYrleuGlnGlyCysGlnArgAspValleuAnlyAspLeuMet 180
QY 541 CACGAGATGGAGTGGTATGATGTTAAATGCGAGCTATACCTGTCMAAGCCTGACTTT 600
DB 181 GlnGlnAnGlnlyleuGlyYrValleuAnAlaSerAnthrCysProlyspRobpPhe 200
QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGATGACAGCTTTGTGAAATTTTG 660
DB 201 IleProGlnSerHispheLeuArgValProvalAsnAspSerPheCysGlnlyIleLeu 220
QY 661 CCGTGTGGACAAATCAGTACATTTTCATGAGAAAGCAAAACCTTCAATGATGTT 720
DB 221 ProTrpLeuAspLyserValaAspHeileGlnlyAlaIysAlaSeranglyCysVal 240
QY 721 CTAGTCACATGTTTACTGCGGATCTCCGCTCCGACCAACATGCTATGCGCTCATCATG 780
DB 241 LeuIleHisCysleuAlaglyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
QY 781 AAGAGATGAGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGA----- 828
DB 261 LysArgMetAspMetSerleuAspGlnAlaTyrArgArgGlnlyspProThrAspArgAla 280
QY 829 -----AAAAGCTACTATATTCCTCAACTTCATTTTTCGGGC 867
DB 281 AspSerArgArgThrGlyMetlysvArgAlaProleuLyserSerleuAn-AlaGlnAl 300
QY 868 CAACCTCGTACATGAGAAAGATTAAAGAACAGA 904
DB 300 AlaIlysrTrpAsnleuGlnArgAlaLeuCyAspArg 312

RESULT 6
O8NST1 PRELIMINARY; PRT; 616 AA.
ID O8NST1;
AC O8NST1;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Lung;
RC Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031643; AAB31643.1; -.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
```

DR PROSITE; PS50056; TYR PHOSPHATASE 2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hypothetical protein 1.
 FT NON TER
 SQ SEQUENCE 616 AA; 67636 MW; 2CB0B14482F2AD72 CRC64;

Alignment Scores:

Pred. No.: 4,67e-126 Length: 616
 Score: 1312.00 Matches: 251
 Percent Similarity: 99.60% Conservative: 1
 Best Local Similarity: 99.21% Mismatches: 1
 Query Match: 80.64% Indels: 0
 DB: 4 Gaps: 0

US-10-029-345a-108_COPY_538_1443 (1-906) x Q99MG5 (1-616)

```

QY 148 TCCTCAAGCTTATTAAGCGAAGTTGCAACAGACAAAGTGTAAATTAACAGAGCTCATC 207
DB 1 CysSerLysLeuMetLysArgArgLeuGlnGlnAspLysValLeuIleThrGluLeuIle 20
QY 208 CAGCATTCAGCGAACAATTAAGTTGACATTGATTCAGATTCAGAGAGTTGTTAGCAT 267
DB 21 GlnHisSerAlaLysHisLysValAspIleAspCysSerGlnLysValValValLysAsp 40
QY 268 CAAAGCTCCCAAGATGTTGCTCTCTCTCTGACATGCTTTTCTACCTGATCTTGGT 327
DB 41 GlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrValLeuGly 60
QY 328 AAACGTGAGAGAGCTTCAACTCTGTTCACTGCTTGACAGGTGGGTTGCTGAGTTCTCT 387
DB 61 LysLeuGlnLysSerPheAsnSerValHisLeuLeuAlaGlyLysPheAlaGluPheSer 80
QY 388 CGTTGTTCCCTGGCCCTCTGTGAAGAAATTCACCTAGTCCCTACCTGATTCCTGAC 447
DB 81 ArgCysPheProGlyLeuCysGlnGlyLysSerThrLeuValProThrCysLysSerGln 100
QY 448 CCTTGCTTACCTGTTGCCAATGTTGGCCCAACCGAATTTCTCCCAATCTTATCTGGC 507
DB 101 ProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuIleLeuGly 120
QY 508 TGCCAGCGAGATGTCCTCAACAAAGAGCTGATCAGACAGATGGAGTTGTTATGTGTTA 567
DB 121 CysGlnAspAspValLeuAsnLysGluLeuMetGlnGlnAsnGlyIleGlyTyrValLeu 140
QY 568 AATGCCAGCTAATCCTGTCGAAGCCTGACCTTATATCCGAGCTTCATTTCTGCGTGG 627
DB 141 AsnAlaSerAsnThrCysProLysProAspPheIleProGluSerHisPheLeuArgVal 160
QY 628 CCTGTGAATGACAGCTTTTGTGAAGAAATTTGCCGTGTTGACCAATCAATGATGATTC 687
DB 161 ProValAsnAspSerPheCysGlnLysIleLeuProIleAspLysSerValAspPhe 180
QY 688 ATTGAAGAAAGCAAAAGCCTCAATGATGTGTTCTAGTGCATGTTTAACTGGATCTCC 747
DB 181 IleGluLysAlaLysAlaSerAsnLysCysValLeuValHisCysLeuAlaGlyIleSer 200
QY 748 CGCTCCGACCATGCTATGCGCTCAATCATCATGAAGATGACATGTTTAAAGTAA 807
DB 201 ArgSerAlaThrIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 220
QY 808 GCTTCAAGATTTGTGAAGAAAGAAAGAACTACTATATCTCCAAATCTTCAATTTCTGGGC 867
DB 221 AlaTyrArgPheValLysGluLysArgProThrIleSerProAsnPheAsnPheLeuGly 240
QY 868 CAACCTCGACTGACTGAGAGAGATTAAGAACCACT 906
DB 241 GlnLeuLeuAspLysGlnLysLysIleLysAsnGlnThr 253

```

RESULT 7

Q99MG5 PRELIMINARY; PRT; 622 AA.
 AC Q99MG5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Map kinase phosphatase-M A2 isoform.
GN DUSP16 OR 3830417M17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RA Yoshikai Y.;
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages."
RL Mol. Cell. Biol. 20:6999-7009 (2001).
DR EMBL; AF345952; AAK35053.1; -.
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IDA.
DR GO; GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro; IPR000340; P:inactivation of MAPK; IDA.
DR InterPro; IPR008343; MAPK phosphatase.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANASE 3; 1.
DR PROSITE; PS00383; TYR PHOSPHATASE 1; 1.
DR PROSITE; PS50056; TYR PHOSPHATASE 2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SQ SEQUENCE 622 AA; 68672 MW; 6C0CA4B6E909B98 CRC64;

```

Alignment Scores:

Pred. No.: 1.1e-111 Length: 622
 Score: 1172.50 Matches: 234
 Percent Similarity: 80.13% Conservative: 8
 Best Local Similarity: 77.48% Mismatches: 5
 Query Match: 72.07% Indels: 55
 DB: 11 Gaps: 1

US-10-029-345a-108_COPY_538_1443 (1-906) x Q99MG5 (1-622)

```

QY 1 ATGCCCATGAGATGATTTGAACTCAATTTGTTACTGAGAGTTGGTGGCTGTGAGAA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrIleSerLeuValAlaLeuGln 20
QY 61 AGTGAACGGAAGAAAGTCTCTTAATGATAGCCGCGCATTTGTGGAATATCAATCC 120
DB 21 serGlyThrGlnLysValLeuLeuIleAspSerArgProPheAlaGluLysAsnThrSer 40
QY 121 CACATTTTGAAGCCATTAAATATCAACTGCTCCAACTTATGAAGGAAAGTTGCAAG 180
DB 41 HisIleLeuGlnAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 200
QY 181 GACAAAGTGTAAATTAACAGAGCTCACTCCAGATTCAGCGAAACATPAAGTTGACATGAT 240
DB 61 AspArgValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY 241 TGCACTCAGAGAGTTGATGATTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
DB 81 CysAsnGlnLysValValValValLysAspGlnSerSerGlnAspValGlySerLeuSerSer 100
QY 301 GACTGTTTCTCACTGATCTTGGGTTAACTGAGAGAGCTTCACTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnArgSerPheAsnSerValHisLeu 120

```

```

QY 361 CTTCAGAGTGGGTTTGCTGAGATCTCTCGTTGTTCCCTGGCCTGTGAGAAATCC 420
    |||||
Db 121 LeuAla----- 122
QY 421 ACTTACTCCCTACTGCAATTTCTACGCCCTTGCTTACTGTGCAACATTGGCCAAACC 480
    |||||
Db 122 ----- 122
QY 481 CGAATTCCTCCCAATCTTTATCTTGCTGCGCAGAGAGATGTCTCAACAAGAGCTGATA 540
    |||||
Db 123 -----AspLeuMet 125
QY 541 CACGAGATGGAGTTGGTATGTTATGTTAAATGACGATACCTGTCCAAAGCCTGACTTT 600
    |||||
Db 126 GlnGlnAnsnGlyIleGlyIleValLeuAnsnAlaSerAnthrCyProlyspProAspPhe 145
    |||||
QY 601 ATCCCCGAGTCTCATTTCTGCGGTGCTGTGATGATGACAGCTTTGTGAGAAATTTTG 660
    |||||
Db 146 IleProGlnSerHisPheLeuArgValProValAnsnSerPheCyGlnIlyIleLeu 165
    |||||
QY 661 CCGTGGTGGACAAATCATAGTATGATTTTATTTGAGAAAGCAAAAGCCTCAATGATGTT 720
    |||||
Db 166 ProTrrPheAspIysSerValAspPheIleGlnIlyAlaIlyAlaSerAnsnGlyCyVal 185
    |||||
QY 721 CTAGTGCACTGTAGCTGGGATCTCCGCTCCGCGCACATGCTATGCGCTACATCATG 780
    |||||
Db 186 LeuIleHisCyLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIlyIleMet 205
    |||||
QY 781 AAGAGATGACATGTCTTTAGATGAGATCTTACAGATTTGTGAAAGAAAGAAAGCTTACT 840
    |||||
Db 206 LysArgMetAspMetSerLeuAspGlnAlaIlyArgPheValIlyGlnIlyArgProThr 225
    |||||
QY 841 ATATCTCCAACATTTCAATTTCTGGGCAACTCTGACATATGAGAAAGATTAAGAAC 900
    |||||
Db 226 IleSerProAnPheAnPheMetGlyGlnLeuMetAspIlyGlnIlySerHisLeuAnsn 245
    |||||
QY 901 CAGACT 906
    |||||
Db 246 GlnThr 247

```

```

DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR HydroLase; kinase.
DR SEQUENCE 300 AA; 33624 MW; 4C61846ACDF0F456 CRC64;

Alignment Scores:
Pred. No.: 1,09e-97 Length: 300
Score: 1036.00 Matches: 217
Percent Similarity: 72.20% Conservative: 9
Best Local Similarity: 69.33% Mismatches: 20
Query Match: 63.68% Indels: 67
DB: 11 Gaps: 2

US-10-029-345a-108_COPY_538_1443 (1-906) x Q9AG15 (1-300)

QY 1 ATGGCCATGAGATGATGGAACTCAATTTGTTACTGAGAGTTGGTCTGTGGA 60
    |||||
Db 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnSerLeuValAlaLeuGln 20
    |||||
QY 61 AGTGAACGGAAAGTGGCTTAATTTGATAGCCGGCCATTTGTGAAATACATATCC 120
    |||||
Db 21 SerGlyThrGlnIlyValLeuLeuIleAspSerArgProPheValGlnIlyArgThrSer 40
    |||||
QY 121 CACATTTTGAAGCCATTAATATCACTGCTCCAAAGCTTATGAGAGAGTTGCAACAG 180
    |||||
Db 41 HisIleLeuGlnAlaIleAnsnIleAnsnIleAnsnIleAnsnIleAnsnIleAnsn 60
    |||||
QY 181 GACAAAGTTAAATTAACAAGCTCATCAGACATTCAGCGAAACATTAAGTTGACATTGAT 240
    |||||
Db 61 AspArgValLeuIleThrGlnLeuIleGlnHisSerAlaIlyHisIlyValAspIleAsp 80
    |||||
QY 241 TGCAGTCAGAGGTTGTAGTTTATCGATCAAGTCCCAAGATGTGCTCTCTCTCA 300
    |||||
Db 81 CysAnsnGlnIlyValIlyValIlyValIlyValIlyValIlyValIlyValIlyValIlyVal 100
    |||||
QY 301 GACTGTTTCTCACTACTTCTGAGTAACTGAGAAAGCTTCAACTGTTCACCTG 360
    |||||
Db 101 AspCyPheLeuThrValLeuLeuGlyIlyBleuGlnArgSerPheAnsnSerValHisLeu 120
    |||||
QY 361 CTTCAGAGTGGGTTGCTGAGTCTCTCTGTTGTTCCCTGGCCTGTGTGAGAAATTC 420
    |||||
Db 121 LeuAla----- 122
QY 421 ACTTACTCCCTACTGCAATTTCTACGCCCTTGCTTACCTGTGCAACATTGGGCCAACC 480
    |||||
Db 122 ----- 122
QY 481 CGAATTCCTCCCAATCTTTATCTTGCTGCGCAGAGAGATGTCTCAACAAGAGCTGATA 540
    |||||
Db 123 -----AspLeuMet 125
QY 541 CACGAGATGGAGTTGGTATGTTATGTTAAATGACGATACCTGTCCAAAGCCTGACTTT 600
    |||||
Db 126 GlnGlnAnsnGlyIleGlyIleValLeuAnsnAlaSerAnthrCyProlyspProAspPhe 145
    |||||
QY 601 ATCCCCGAGTCTCATTTCTGCGGTGCTGTGATGATGACAGCTTTGTGAGAAATTTTG 660
    |||||
Db 146 IleProGlnSerHisPheLeuArgValProValAnsnSerPheCyGlnIlyIleLeu 165
    |||||
QY 661 CCGTGGTGGACAAATCATAGTATGATTTTATTTGAGAAAGCAAAAGCCTCAATGATGTT 720
    |||||
Db 166 ProTrrPheAspIysSerValAspPheIleGlnIlyAlaIlyAlaSerAnsnGlyCyVal 185
    |||||
QY 721 CTAGTGCACTGTAGCTGGGATCTCCGCTCCGCGCACATGCTATGCGCTACATCATG 780
    |||||
Db 186 LeuIleHisCyLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIlyIleMet 205
    |||||

```

QY 781 AAGAGATGACATGCTTATGATGAGCTTACAGATTGTAAGAA----- 828
 Db |||||
 Db 206 LyaragwetraspmettererleuaspglualyrgarglnlyserProthInaspargala 225
 QY 829 -----AAAAGCTTACTATATCTCCAAACTTCATTTTGGGC 867
 Db 226 AspserArgrThrclymelysarGlaProleuyserserleuasn-AlaGlual 245
 QY 868 CAATCTGACTGATGACGAAAGATTAAAGAACGGA 904
 Db |||||
 Db 245 aaAlalystrpenleuGluarGlaLeuCySarGarg 257

RESULT 9

Q86S8 PRELIMINARY; PRT; 625 AA.
 ID 086S8
 AC 086S8
 DT 01-JUN-2003 (TRENBLREL. 24, Created)
 DT 01-JUN-2003 (TRENBLREL. 24, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE Similar to dual specificity phosphatase 8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 RN NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC STRSUB=Brain, and Astrocytoma;
 RA Strausberg R.;
 RL EMBL/BC045110; JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR GO; GO:0017017; F.MAP kinase phosphatase activity; IEA.
 DR GO; GO:0006470; P.protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR002965; P-rich exten.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR00387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR PRINTS; PR01217; PRICHEXTENSN.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANES_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 SQ SEQUENCE 625 AA; 65826 MW; C7C808407B724FFC CRC64;

Alignment Scores:

Pred. No.: 8.84e-93 Length: 625
 Score: 989.00 Matches: 190
 Percent Similarity: 81.40% Conservative: 55
 Best Local Similarity: 63.12% Mismatches: 54
 Query Match: 60.79% Indels: 2
 DB: 4 Gaps: 2

US-10-029-345a-108_COPY_538_1443 (1-906) x Q86S8 (1-625)

QY 1 ATGGCCCATGAGATATGGAATCTAAATGTT---ACTGAGAGTTGGTGGCTGCTG 57
 Db |||||
 Db 1 MetAlaGlyAspArgLeuProArgLyValMetCasPalalysylsleuAlaSerleu 20
 QY 58 GAAAGTGGACGGAAGAAAGTGGCTAATGATGACCGGCACTTTGTGGAATACATACA 117
 Db |||||
 Db 21 ArgGlyLyProGlyLyProLeuValIleAspSerArgrSerPheValGluTyrrhanser 40
 QY 118 TCCGACATTTGGAAGCATTAATATCACTGCTCCAGACTTAAGACGAAAGTTGCA 177
 Db |||||
 Db 41 TrpHisValIleuSerValIleuIleCysCysSerLySleValLySarGargLeuGln 60
 QY 178 CAGACAAAGTGTAAATTCAGAGCTCACCAGCATTCAGCGAAACATTAAGTTGACATT 237
 Db |||||
 Db 61 GlnGlyLyValThrIleAlaGluIleuIleGlnProAlaIlaArgSerGlnValGluAla 80

QY 238 GATTGACATGACAGAGTTGTAGTTATGATACATCAAGAGCTCCAGAGTTGGCTCTCT 297
 Db |||||
 Db 81 ThrGluProGlnAspValValValLyZaspGlnSerThrArGaspAlaSerValIleAla 100
 QY 298 TCAGACTGTTTTCTGACTGTACTTCTGGGTAACTGAGAAAGAGCTTCAACTGTGTCAC 357
 Db |||||
 Db 101 AlaAspSerPheLeuSerIleuLeuSerLyLeuAspGlyCysPheAspSerValAla 120
 QY 358 CTGCTTGACAGGAGGTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTAAGAGAA 417
 Db |||||
 Db 121 IleLeuThrGlyLyGlyPheAlaThrPheSerSerCysPheProGlyLeuGlySerGlu 140
 QY 418 ---TCCACTTAGTCCCTACCTGATTTCTCAGCTTCTGCTTCTGCTTCCAAATTTGGG 474
 Db |||||
 Db 141 ProAlaIlaLeuLeuProMetSerLeuSerGlnProCysLeuProValProSerValGly 160
 QY 475 CCAACCCGAATTTCTCCCAATCTTTATCTTGCTGCCAGAGATGCTCAACAGAG 534
 Db |||||
 Db 161 LeuThrArgIleuProHisLeuTyrrLeuGlySerGlnLyAspValLeuAsnLyAsp 180
 QY 535 CTGATACAGACAGATGGATGGATTGTTATGTTAAATGACGATATACCTGTCCAAAGCTT 594
 Db |||||
 Db 181 LeuLeuThrGlnAsnGlyIleSerTyrrValLeuAsnAlaSerAsnSerCysProLyPro 200
 QY 595 GACTTTATCCCGAGCTCATTTCTGCGCTGCTGCTGTAATGACAGCTTTGTGAGAA 654
 Db |||||
 Db 201 AspPheIleCysGluSerIleArgPheMetArgValProIleAsnAspAsnTyrrCysGlu 220
 QY 655 ATTTGCGCTGCTGTTGACAAATCACTGATTTCAATGATTTGAGAAAGAAAGCTTCATGCA 714
 Db |||||
 Db 221 LeuLeuProTyrPheAspLySerIleGluPheIleAspLyAlaLyLeuSerCys 240
 QY 715 TGTGTTTACGTCACTGTTAGCTGGATCTCCCGCTCCGCCCACTGCTATGCTTAC 774
 Db |||||
 Db 241 GlnValIleValHisCysLeuAlaGlyIleSerArgrSerValThrIleAlaIleAlaTy 260
 QY 775 ATCAAGAGAGAGATGACATGCTTTAGATGATGATGATGATGATGATGATGATGATGAT 834
 Db |||||
 Db 261 IleLeuLyThrMetGlyMetSerSerAspAspAlaTyrrArgPheValLyAspArgArg 280
 QY 835 CCTACTATATCTCCAAACTTCATTTTCTGGGCAACTCTGACTATGAGAAAGATT 894
 Db |||||
 Db 281 ProSerIleSerProAsnPheAsnPheLeuGlyGlnLeuGluTyrrGluArgSerLeu 300
 QY 895 AAG 897
 Db |||||
 Db 301 Lys 301

RESULT 10

Q77S29 PRELIMINARY; PRT; 665 AA.
 ID 077S29
 AC 077S29
 DT 01-OCT-2003 (TRENBLREL. 25, Created)
 DT 01-OCT-2003 (TRENBLREL. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE Dusp8 protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Kleusner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Hopkins R.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Ushin T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

[illegible]

Db |||||
Qy 21 SerGlyInrGlutylValLeuLeuIleAspSerArgProPheValGluTrpAsnThrSer 40
Db 121 CACATTTTGGAAAGCCATTATATCACTGCCTCAGCGCTTAAGAAGCGAAGGTTCGAACAG 180
Qy 41 HstIleuEnGIuaIalIeAnIIeaNCysSerLysIueWtLyBAAGAglEuGIln 60
Db 181 GACAAGNGTAAATTCACAGAGCTCATCCAGCATTTGACGGAAMAATAAGTTGACATTGAT 240
Qy 61 AspYrsValLeuIleThrgIuLeuIlegInHsserAlayShIslysValaplleasp 80
Db 241 TGCAGTCAGAAAAGGTTTGATTTTAGCATCAAAGCTCCCAGAGTGTGGCTCTCTCTTCA 300
Qy 81 CysserdInlyrsValIvalIValtyrAsplInserserdInspAlalaserLeuserdSer 100
Db 301 GACTGTTTTCTCACGTGTCCTCTGGGTAACTGGAGAAGAAGCTTCAACTCTGTACCCTG 360
Qy 101 AspCySPheLeuThrValLeuenuGLyLysleuGIuLysSerPhesAnserrValHsstieu 120
Db 361 CTTCGACGGTGGGTTTGCTGAGTTCTCTCGTTTTCCTGGCGCTCTGTGAAGAAAAATCC 420
Qy 121 LeuAlaagLY--Ala----- 124
Db 421 ACTCTAGTCCTCAGCTGCATTTCTCAGCGCTTGCTTACCTGTTGCCAACATTGGCCAACC 480
Qy 124 ----- 124
Db 481 CGAATTCCTCCACATCTTTATCTTGSGCTCCAGAGATGTCCTCAACAAGAGCTGATA 540
Qy 125 ----- 126
Db 541 CAGCAGATGGATGGATTGATGTTATGTTAAATGCAGCACTAAGCTTCCAAAGCC 593
Qy 126 laalaGIurtpasptRpleucysVallysCysgInGIntyrLeuSerLysAla 143

RESULT 13

ID Q8ST19 PRELIMINARY; PRT; 657 AA.

AC Q8ST19;
DT 01-JUN-2002 (TEMBREL. 21, Created)
DT 01-JUN-2002 (TEMBREL. 21, Last sequence update)
DT 01-OCT-2003 (TEMBREL. 25, Last annotation update)
DE Hypothetical protein F08B1.1a.
GN F08B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodermiae; Caenorhabditis.
OX NCBI_TaxID=6239;

RM [1]_TaxID=6239;
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol NZ;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol NZ;
RA Chissos S.;
RT "The sequence of C. elegans cosmid F08B1."
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol NZ;
RA Waterston R.;
RT "Direct Submissions".
RL EMBL/U3178; AKS68299.1; -.
DR HSSP: Q16828; IMKP.
DR MorphP: F08B1.1a; CE27918.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.


```

QY 43 TTGGTCTCTGCTGAAAAGTGAAGAAAAGTGGCTTAATGATTCAGCCGCAATT 102
Db |||||
15 LeuLysArgLeuMetLysAspArgLysAlaLysCysLeuLeuLeuAspCysArgSerPhe 34
QY 103 GTGGAATACATATACATCCCATTTTGGAGCCCAATTAATATCAATGCTCCAGCTTAG 162
Db |||||
35 LeuAlaPheSerAlaGlyHisLeuArgGlyAlaValAlaSnIleArgCysSerThrIleVal 54
QY 163 AAGCCAGG-----TTGCAACAGCAAGAGTGTATTTACAGACTCATCCAG 210
Db |||||
55 ArgAspArgAlaLysGlySerValSerLeuAspGlnIleLeuSerGlyAsp----- 71
QY 211 CATTACAGGAAACATAGTTGATGATTCAGATGCACAGAGGTTGTATGATTCAGATCA 270
Db |||||
72 AspGlnAlaArgSerArgLeuLysSerGlyLeuTyrSerAlaValIleLeuTyrAspIle 91
QY 271 AGCTCCCAAGATGTTGCTCTCTCTCTTCAGACTGTTTCTACAGTACTCTGGATAA 330
Db |||||
92 ArgSerSerAspThrAsnThrMetLysAspAspSerThrIleThrLeuValHisAsnAla 111
QY 331 CTG---GAGAAGAGCTTCAACTCT---GTTCACCTGCTTGAAGGTGGTGTGATGCTG 384
Db |||||
112 LeuCysArgAspThrPheSerThrGlnValTyrIleLeuLysGlyGlyTyrAspArgPhe 131
QY 385 TCTCGTTGTTCCCTGCTCTCTG-----GAAAGAAA 417
Db |||||
132 SerThrGlnTyrProAspTyrCysLeuLysThrArgThrLeuSerValSerSerSerGln 151
QY 418 TCCACTCATCCCTACCTGATTTCTCAGCCTTCTTACCTGTTGCAACATTTGGCGCA 477
Db |||||
152 SerSerMetGlnSerSerCysLeuSerValAlaThrProGlnHisAspGlnGlyPro 171
QY 478 ACCCAATTTCTCCCAATCTTTATCTTGCTGCTGACAGAGATGCTTCAACAGAGCTG 537
Db |||||
172 ValGlnIleLeuProPheLeuPheLeuGlySerAlaLeuHisAlaSerLysAspMet 191
QY 538 ATACAGCAGAAATGAGATTTGTTATGTTGTTAAAGCCATTAACCTGTCACAAAGCCTGAC 597
Db |||||
192 LeuAspArgMetGlyIleSerAlaLeuLeuAsnValSerSerAsnCysProAsn---His 210
QY 598 TTTATCCCAAGCTCATTTCTCGCGTGCCTGATGATGATGATTTTGGAGAAATT 657
Db |||||
211 PheGlnIleLysAspTyrGlnTyrLysCysIleProValGlnAspAsnHisLysGlnAspIle 230
QY 658 TTGCGCGTGTGGAACAATCAGTATGATTTGATGAGAAAGCAAGCCTCCATGATGAT 717
Db |||||
231 SerSerThrPheIleGlnAlaIleGlnPheIleAspSerValLysAspSerAsnGlyArg 250
QY 718 GTTCTGTCGACATGTTTACCTGAGATCTCCGCTCCGACACATCGGATATGCTGCTATC 777
Db |||||
251 ValLeuValHisCysGlnAlaGlyIleSerArgSerAlaThrIleCysLeuAlaTyrLeu 270
QY 778 ATGACAGAGATGACATGCTCTTATGATAGATTTACAGATTTGTAAGAAAAGAAAGCT 837
Db |||||
271 MetLysLysLysArgValArgLeuGlnIleValPheGlnPheValLysGlnArgSer 290
QY 838 ACTATATCTCCCAACTTCAATTTTCTGGGCAACTCTCTGATGATGATGAGAAAGAT 894
Db |||||
291 IleIleSerProAsnPheSerPheMetGlyGlnLeuLeuGlnPheGlnSerGlnVal 309
RESULT 15
Q8R3L3
ID Q8R3L3 PRELIMINARY, PRT, 483 AA.
AC Q8R3L3;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DE 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Dual specificity phosphatase 10.
GN DUSP10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RA SEQUENCE FROM N.A.
RP Strauberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RA SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J, and NOD;
RC TISSUE=Pancreas, Thymus, and Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
the KIRKS Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002)."
DR EMBL; BC025066; AAH25066.1; -
DR EMBL; AK035293; BAC29019.1; -
DR EMBL; AK050528; BAC34308.1; -
DR EMBL; AK088024; BAC40102.1; -
DR EMBL; AK088186; BAC40196.1; -
DR EMBL; AK088357; BAC40300.1; -
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1927070; Dusp10.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00782; DSPC; 1. _phosphatase.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS0206; RHODANES 3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE 1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE 2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase.
SQ SEQUENCE 483 AA; 52532 MW; 7797A1877D986A8 CRC64;
Alignment Scores:
Pred. No.: 3,12e-39
Score: 469.00 Length: 483
Percent Similarity: 58.05% Matches: 106
Best Local Similarity: 35.57% Conservative: 67
Query Match: 28.83% Mismatches: 95
DB: 11 Indels: 30
Gaps: 7
US-10-029-345a-108_copy_538_1443 (1-906) x Q8R3L3 (1-483)
QY 79 CTGCTAATTGATGATCCGCGCATTTGTGGAATACATCAATCCCATTTTGAAGCCATT 138
Db |||||
174 ValIleIleAspCysArgProPheMetGlnTyrAsnLysSerHisIleGlnGlyAlaVal 193
QY 139 AATATCACTGCTCC---AAGCTTATGAAGCAAGTTGCAACAGCAAGATGTTAATT 195
Db |||||
194 HisIleAsnCysAlaAspLysIleSerArgArgLeuGlnGlnGlnGlnGlyIleThrVal 213
QY 196 ACAGAGCTCATCCAG---CATTCAGCGAAACATPAGGTTACATTTGATTCAGTCAAG 252
Db |||||
214 LeuAspLeuIleSerCysArgGlnGlyLysAspSerPheLysArgIlePheSerLysGln 233
QY 253 GTTAGTTTATGATGATCAAGCTCCCAAGCTGCTCTCTCTTCAAGATGTTTCTC 312
Db |||||
234 IleIleValTyrAspGlnValAsnThrAsnGlnProSerArgValThrProSerGlnProLeu 253
QY 313 ACTGACTTCTGGGTAACTGAGAAAGAGCTTCAACTGTTGCTACCTGTTGCAAGTGG 372
Db |||||
254 HisIleValLeuGlnSerLeuLysArgGlnGlyLysGlnProLeuValLeuLysGly 273
QY 373 TTGCTGAGTTCTCTGTTGTTTCCCTGCGCTCTGTGAAGAAA----- 417

```

```
Db 274 LeuSerSerPheIysGlnAsnHisGlyAsnLeuCysAspAsnSerLeuGlnLeuGlnIu 293
QY 418 -----TCCACTCTAGTCCCTTACCTGATTTCT 444
Db 294 CysArgGlnValGlyGlyAlaSerAlaIAserSerMetLeuProGln-----Ser 311
QY 445 CAGCCTTGCTTACCT---GTTGCCACATTTGGGCCAACCCGAATTTCTCCCAATCTTTAT 501
Db 312 ValProThrThrProAspIleGlnAsnAlaGlnLeuThrProIleLeuProPheLeuPhe 331
QY 502 CTGGCTGCGCAGCGAGATGTCTTCACACAGAGCTGTATACAGCAATGGGANTGTTAT 561
Db 332 LeuGlnAsnGlnAspAlaGlnAspLeuAspThrMetGlnArgLeuAsnIleGlyTyr 351
QY 562 GTGTAAATGCCAGC-----TATACCTGTCCAAAGCCTGACTTTATCCC 606
Db 352 ValIleAsnValThrThrHisLeuProLeuTyrHisTyrGlnIlysgIlyLeuPhe----- 369
QY 607 GAGTCTCATTTCTGCGTGTGCTGTGATGACAGCTTTTGAGAAATTTTGCGGTGG 666
Db 370 -----AsnTyrIlysalArgLeuProAlaThrAspSerAsnIlysgIlnAsnLeuArgGlnTyr 387
QY 667 TTGGACAATCACTAGATTTCAATTGAGAAACAAAGCTTCAATGGATGTGTTCTAGTG 726
Db 388 PheGlnGlnIuAlaPheGlnIuPheIleGlnIuAlaHisGlnCysGlysgIlyLeuLeuIle 407
QY 727 CACTGTTTACCTGGGATCTCCGCTCCGCCACCATCGCTATCGCTTACATCATGAAGAG 786
Db 408 HisCysGlnAlaGlyValSerArgSerAlaThrIleValIleAlaTyrLeuMetIyHis 427
QY 787 ATGGACATGTCTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGACCTACTATATCT 846
Db 428 ThrArgMetThrMetThrAspAlaTyrIlyPheValIySalArgProIleIleSer 447
QY 847 CCAAACTTCAATTTCTGGGCCCAACTCTGACTATGAGAGAGATTAAAGAC 900
Db 448 ProAsnLeuAsnPheMetGlyGlnLeuLeuGlnPheGlnIuAspLeuAsnAsn 465
```

Search completed: June 21, 2004, 12:40:25
Job time : 53.8459 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_n2p model

Run on: June 21, 2004, 12:28:21 ; Search time 43.4105 Seconds
(without alignments)
11793.817 Million cell updates/sec

Title: US-10-029-345a-108_COPY_538_1443

Perfect score: 1627
Sequence: 1 atggccacatgagatgcttgg.....agaagatgaaccagact 906

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlp
-Q=/cgr2.1/USPTO.spool.p/US10029345/runat.21062004.122814.4167/app.query.fasta_1.3278
-DB=A.Geneseq.29Jan04 -QFMT=fastaan -SUFFIX=rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345.QCEN.1.1.128 -runat.21062004.122814.4167 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELACK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq.29Jan04.*

1: geneseqp1980a.*
2: geneseqp1980a.*
3: geneseqp2000a.*
4: geneseqp2001a.*
5: geneseqp2002a.*
6: geneseqp2003a.*
7: geneseqp2003bs.*
8: geneseqp2004a.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1564	96.1	665	5	ABR52381 Protein r
2	1564	96.1	665	5	ABR52407 Protein r
3	1552	95.4	665	4	AAE04834 Human SGP
4	1552	95.4	665	4	AAU09016 Human dua
5	1552	95.4	665	5	AAU09156 Human dua
6	1552	95.4	665	5	AAU09946 Protein B
7	1552	95.4	665	5	AAU75789 Human pro
8	1552	95.4	665	5	ABR97946 Human pro
9	1552	95.4	665	5	AAU79929 Human dua
10	1552	95.4	665	5	ABR97291 Novel hum

11	1552	95.4	666	4	ABR20325	ABR20325 Human pro
12	1545	95.0	665	5	ADA54744	ADA54744 Human pro
13	1544	94.9	665	5	AAU79161	AAU79161 Human dua
14	1544	94.9	665	5	ABR52352	ABR52352 Protein r
15	1544	94.9	672	4	AAU25744	AAU25744 Human pro
16	1542	94.8	665	5	AAU79162	AAU79162 Human dua
17	1537	94.5	302	5	ABR52425	ABR52425 Protein r
18	1526.5	93.8	664	5	ABR52424	ABR52424 Protein r
19	1482	91.7	660	5	ABR52385	ABR52385 Protein r
20	988	60.7	625	5	ABR52382	ABR52382 Protein r
21	988	60.7	625	5	ABR52350	ABR52350 Protein r
22	988	60.7	625	6	ABR73440	ABR73440 Human dua
23	985	60.5	663	5	AAW29150	AAW29150 Dual -spec
24	985	60.5	663	5	ABR52351	ABR52351 Protein r
25	742.5	45.6	375	4	ABG07902	ABG07902 Novel hum
26	741	45.5	140	5	ABR52404	ABR52404 Peptide r
27	740	45.5	117	5	AAU79159	AAU79159 Human dua
28	716	44.0	140	5	ABR52405	ABR52405 Peptide r
29	715.5	44.0	579	7	ADBE08458	ADBE08458 Novel pro
30	654.5	40.2	170	4	ABR66436	ABR66436 Human MAP
31	654.5	40.2	170	4	AAE06780	AAE06780 Human pro
32	476	29.3	482	4	ABR73225	ABR73225 Human pho
33	469	28.8	444	5	AAO20515	AAO20515 Protein o
34	469	28.8	482	3	AAI18655	AAI18655 A human r
35	469	28.8	482	3	ABR29641	ABR29641 Human dua
36	469	28.8	482	5	ABR52384	ABR52384 Protein r
37	469	28.8	482	5	ABJ05600	ABJ05600 Breast ca
38	467.5	28.7	381	5	ABR52383	ABR52383 Protein r
39	467.5	28.7	381	6	ABP96803	ABP96803 Human COP
40	462.5	28.4	381	6	ABP57087	ABP57087 Mouse MKP
41	461.5	28.4	381	7	ADD48300	ADD48300 Rat Prote
42	461.5	28.4	381	7	ADG62625	ADG62625 Rat Prote
43	440.5	27.1	368	5	AAE14240	AAE14240 Human pro
44	440.5	27.1	419	6	ABP55026	ABP55026 Human dua
45	428	26.3	394	4	AAE67627	AAE67627 Amino aci

ALIGNMENTS

RESULT 1	ABR52381	ABR52381 standard; protein; 665 AA.
ID	ABR52381	
AC	ABR52381	
XX	19-JUN-2003	(first entry)
XX	Protein relating to the invention SEQ ID NO: 109.	
XX	antiapoptotic; hepatotropic; nephrotropic; antiarthritic;	
KW	antiapoptotic; cardiact; cytosolic; gene therapy; liver disease;	
KW	proliferative disorder; renal failure; cardiovascular disorder;	
KW	immunological disorder; arthritis; psoriasis; congenital heart defect;	
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
XX	Homo sapiens.	
OS	WO200257460-A2.	
XX	25-JUL-2002.	
XX	20-DEC-2001; 2001WO-US050459.	
XX	20-DEC-2000; 2000US-025668P.	
PR	30-MAR-2001; 2001US-028018P.	
PR	01-MAY-2001; 2001US-0287735P.	
PR	05-JUN-2001; 2001US-0295848P.	
PR	25-JUN-2001; 2001US-0300465P.	
XX	(BRIM) BRISTOL-MYERS SQUIBB CO.	
PA	Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,	
XX	PI Stiemers N, Bol D, Schieven G, Finger J, Todderud CC, Basolino D;	

PI Krystek S, Mcatee P, Suchard S, Banas D,
 XX WPI; 2002-599721/64.
 DR N-PSDB; ACC60559.
 XX
 PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 XX
 PS Claim 5; Fig 12; 801pp; English.

CC The invention relates to a novel isolated nucleic acid comprising a
 CC polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotides fully defined in the specification. The polynucleotide of
 CC the invention has antiproliferative, hepatotropic, nephrotropic,
 CC antirhectic, antipsoriatic, cardiant, and cytostatic activity. The
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention
 XX
 SO Sequence 665 AA;

Alignment Scores:
 Pred. No.: 1,18e-170 Length: 665
 Score: 1564.00 Matches: 302
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.13% Indels: 0
 DB: 5 Gaps: 0

US-10-029-345A-108_COPY_538_1443 (1-906) x ABR52381 (1-665)

QY 1 ATGGCCCATGAGTATGGAAGTCAATGTTACTGAGAGGTGGCTGCTGCGAA 60
 DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThiGlnArgLeuValAlaLeuLeuIu 20
 QY 61 AGTGAACGGAAGAAAGTGTCTAATGATAGCCGCCCTTTGTGAATACATACATCC 120
 DB 21 SerGlyThrGluLysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 QY 121 CACATTTGGAGACCATTAATATCAACTGCTCCAGCTTAAGAGCGAAAGTTGCAAG 180
 DB 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerIleuMetLysArgAlaGluGln 60
 QY 181 GACAAAGTGTATATACAGAGCTCATCCAGCTTACAGGAAACATTAAGTTGACATTGAT 240
 DB 61 AspLysValLeuIleThiGluLeuIleGlnHisSerIleAlaLysValAlaSerLeu 80
 QY 241 TGCAGTCAGAGGTGTAGTTTACGATCAAGCTCCCAAGATGTGGCTCTCTCTTCA 300
 DB 81 CysSerGlnLysValValValValValValValValValValValValValValValVal 100
 QY 301 GACGTTTTTCTACTGTACTTCTGGGTAACTGGAAGAAAGCTTCAACTCTGTACCTG 360
 DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
 QY 361 CTTCAGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGGCTCTGTAAGAGAAATCC 420
 DB 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGlnLysSer 140
 QY 421 ACTTAAGTCCCTACCTGATTTCTCAGCCCTTGCTTACCTGTTGCAACATTTGGGCAAC 480
 DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 QY 481 CGAATTTTCCCAATCTTATCTTGCTGCGCAAGAGATGTCCTCAACAGAGCTGATA 540
 DB 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgPheValLeuAsnLysGlnLeuIle 180

QY 541 CAGCAGATNGGATTGCTTATGTGTAAATGCGACGTAATACCTGCAAGGCTGACTTT 600
 DB 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerIleThrCysProLysProAspPhe 200
 QY 601 ATCCCGAGTCTCAATTCCTGCGTGTGCTGTAATGACGCTTTGTGAGAAATTTTG 660
 DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
 QY 661 CCGTGTGGAGCAATTCAGTATGATTTCTATGAGAAAGCAAGCTCAATGATGATGTT 720
 DB 221 ProTyrPheAspLysSerValAspPheIleGlnLysAlaLysAsnGlnCysVal 240
 QY 721 CTATGCACTGTTTAACTGCGATCTCCGCTCCGCCCAACATGCTTATGCTTACATCAG 780
 DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 QY 781 AAGAGATGAGCAATCTTTAGATGAGACTTACAGATTTGTGAAAGAAAGAAAGACTTACT 840
 DB 261 LysArgPheCysPheSerLeuAspGlnAlaTyrArgPheValLysGlnLysArgProThr 280
 QY 841 ATATCTCAAACTTCAATTTCTGCGCAACTCTGAGCTATGAGAAAGATTTAAGAAC 900
 DB 281 IleSerProAsnPheAsnPheLeuGlyIleLeuLeuAspTyrGlnLysIleLysAsn 300
 QY 901 CAGACT 906
 DB 301 GlnThr 302

RESULT 2
 ABR52407
 ID ABR52407 standard; protein; 665 AA.
 XX
 AC ABR52407;
 XX
 XX 19-JUN-2003 (first entry)
 DE Protein relating to the invention SEQ ID NO: 148.
 XX
 KW antiproliferative; hepatotropic; nephrotropic; antirhectic;
 KW antipsoriatic; cardiant; cytostatic; gene therapy; liver disease;
 KW proliferative disorder; renal failure; cardiovascular disorder;
 KW immunological disorder; arthritis; psoriasis; congenital heart defect;
 KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
 OS Homo sapiens.
 XX
 PN WO200257460-A2.
 XX
 PD 25-JUN-2002.
 XX
 PF 20-DEC-2001; 2001WO-US050459.
 XX
 PR 20-DEC-2000; 2000US-025686BP.
 PR 30-MAR-2001; 2001US-0280186P.
 PR 01-MAY-2001; 2001US-0287735P.
 PR 05-JUN-2001; 2001US-0295848P.
 PR 25-JUN-2001; 2001US-0300465P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Jackson DG, Feder J, Nelson T, Mintler G, Ramanathan C, Lee L,
 PI Siemers N, Bol D, Schieven G, Finger J, Todderd CG, Basolino D,
 PI Krystek S, Mcatee P, Suchard S, Banas D;
 DR WPI; 2002-599721/64.
 DR N-PSDB; ACC605572.
 XX
 PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 PS Disclosure; Fig 19; 801pp; English.

XX The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antiatherctic, antipositastic, cardiant, and cycostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention

SQ Sequence 665 AA;

Alignment Scores:	
Pred. No.:	1,18e-170
Score:	1564.000
Percent Similarity:	100.000
Best Local Similarity:	100.000
Query Match:	96.13%
DB:	5
Length:	665
Matches:	302
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) X ABR52407 (1-665)

Qy	1	ATGGCCCATGAGTATGATGGAACTCAAAATGTTGTACTGAGAGGTTGGTGGCTCTGCTGAA	60
Db	1	MetAlHhIGLWetIleGlyThrGlnIleValImrGluNagLeuValAlaLeuIleGlu	20
Qy	61	AGTGAACGGAAAAAGCTGCTTAATGTGATAGCCGGCCATTTGTGGAAATCAATACATCC	120
Db	21	SerGlyThrGluNagValLeuLeuIleApsSerArgProPheValGluNagYAsnThrSer	40
Qy	121	CACATTTTGGAAAGCCATTAATATCAACGCTCCAGCTTAATGAAGCGAAGGTGCAACAG	180
Db	41	HsIleLeuGlnIuAlaIleAsnIleAsnCySerIlySleuMetIlyAAsArgLeuGlnGln	60
Qy	181	GACAAAGTGTAAATTACAGAGCTCATCCGACATTCAGGAAACATTAAGTTGACATTGAT	240
Db	61	AspIlyValLeuIleThrGluLeuIleGlnIhAsSerAlaIlyBHisIlyValAspIleAsp	80
Qy	241	TGCAGTCAGAGGTTGTAGTTACGATCAAGACTCCCAAGATGTTGCTCTCTCTTCA	300
Db	81	CysSerGlnIlyValAlaValIlyTrAspGlnSerSerGlnAspValAlaSerLeuSerSer	100
Qy	301	GACTGTTTTCTACCTGCTACTCTCGGGTAAACGAGAAAGAGCTTCACTGTCACCTG	360
Db	101	AspCysPheLeuThrValLeuLeuGlyIlySleuGlyIlySerPheAsnSerValHisIleu	120
Qy	361	CTTGACAGTGGGTTTGCTGAGTTCCTCGTGTGTTCCCTGGCCTCTGTGAAGAAAATCC	420
Db	121	LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuCyGlnGlyIlySer	140
Qy	421	ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTCTCACTGTTGCCAACAATTGGGCCAAC	480
Db	141	ThrLeuValaProThrCysIleSerGlnProCysIleuProValAlaAsnIleGlyProThr	160
Qy	481	CGAATTCCTCCCAATCTTTATCTTGAGCTGCCGACGAGATGTCTTCAACAGAGACTGATA	540
Db	161	ArgIleLeuProAsnLeuIlyIleuGlyCysGlnArgAspValLeuAsnIlyGlnLeuIle	180
Qy	541	CAGCAGAAATGGATTTGGTTATGTGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT	600
Db	181	GlnGlnAsnAlaIlyIleGlyIlyValaLeuAsnAlaSerIlyThrCysProIlyProAspPhe	200
Qy	601	ATCCCCGAGCTCATTTCCGCGCTGGCGCTGGAGATGACAGCTTTTGGAGAAAATTTTG	660
Db	201	IleProGlnSerHisPheLeuArgValaProValaAsnAspSerPheCysGlnIlyIlyIleu	220
Qy	661	CCGTGTTGGCAAAATCAGTAGATTTGATGTAGAAAGCAAAAGCCTCCATGATGATGTT	720

Db	221	Protripeuaplyserservalaspheilegulusalalyalsersamlylcyvsval	240
Qy	721	CTATGCACTGTTAGCTGGAGTCTCCGCTCCGGCCACATCGCTATCGCTACATCAG	780
Db	241	Leuvalhscysleuhalaglylleserxrgseralathrllealallelathrylilemet	260
Qy	781	AAGAGATGACATGTCCTTTAGATGAGACTTACAGATTGTGAAAGAAAAGACTTACT	840
Db	261	LysrsmetclapmetSerleuaspolulalathyrasgphevalylgslulysargProthr	280
Qy	841	ATATCTCCAAACTCAATTTTCTGGGCCCACTCTGACATATAGAGAGAGATTAAAGAC	900
Db	281	IlleserProanpheanPheleuglyglnleuenuaptyrGluulysrlleysaen	300
Qy	901	CAGACT 906	
Db	301	Glnthr 302	

RESULT 3

ID AAE04834 standard; protein; 665 AA.

AC AAE04834 ;

DT 10-SEP-2001 (first entry)

Human SGP002 phosphatase polypeptide.

KM Human; PNP002 phosphatase polypeptide; phosphatase-related disease;
KM immune-related disorder; ocular disease; organ transplant rejection;
KM infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KM metabolic disorder; hematopoietic cancer; myod disorder; cardiant;
KM Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KM cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KM attention disorder; cognition disorder; psychotic disorder; cytoskeletal;
KM neurological disorder; vinculin; nootropic; cerebroprotective; therapy;
KM neuroprotective; antibacterial; vulnerability; transilliser; antislammatic
KM hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KM antitumoral; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KM MKP; migraine; chromosome 12p11.1;p12.1.

OS Homo sapiens

FH	Key	Location/Qualifiers
153		

```
FT      /label= catalytic_domain
DE      150 307
```

/label= phosphatase_domain

PN MO200146394-A2
 YX

28-JUN-2001
PD
XX

PF 21-DEC-2000; 2000MCU-03034/36
XX

PR	28-DEC-1999;	99US-0175766P
EN	21-DEC-1999;	99US-0175766P

PR 31-JAN-2000; 2000US-0179301P.

PA (SUGE-) SUGEN INC.

PI PLOWMAN GD, MARTINEZ R, WHY

XX

DR N-PSDB; AAD09492.

Novel phosphatase polypeptide useful for treating cancers, immune-related disorders and cardiovascular disease, brain or neuronal

XX

CC designated 2117 and 38692, the nucleic acids encoding them (including
CC fragments, allelic variants, their complements or nucleic acids that
CC hybridize to them) and antibodies raised against the proteins. The
CC antibody is useful for detecting the presence of the polypeptide, and the
CC nucleic acid fragments are useful for detecting the presence of the
CC nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
CC antisense sequences) are useful for modulating the activity or expression
CC of the polypeptide/nucleic acid, useful for the treatment of e.g.
CC hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
CC listed in the specification) liver disorders, erythroid associated
CC disorders (e.g. haemolytic anaemia) cellular proliferative or
CC differentiative disorders, leukaemias (e.g. acute myeloid leukaemia),
CC carcinomas, sarcomas, metastatic cancers. Agents modulating 2117 and
CC 38692 are also useful for modulating the proliferation, survival,
CC migration or differentiation of a 38692 or 2117-expressing cell. The
CC polypeptide and nucleic acids are useful for identifying modulating
CC agents. The present sequence represents the dual specificity phosphatase
CC 2117
XX
XX

SQ Sequence 665 AA.

Alignment Scores:

Pred. No.:	2.85e-169	Length:	665
Score:	1552.00	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	0
Query Match:	95.39%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) x AAU09016 (1-665)

QY 1 ATGGCCCATGAGATGATGGAATGTTTACTGAGAGGTTGGCTCTGCTGGA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAAGTGTGTAATGATGAGCGGCACTTGTGGAATCAATACATCC 120
DB 21 SerGlyThrGluValLeuLeuIleAspSerThrProPheValGluTyrAsnThrSer 40
QY 121 CACATTTGGAGCCATTAAATCACTGCTCCAGGTTTGAAGCGAAGTTGCAACAG 180
DB 41 HisIleLeuGlnIleAlaIleAsnIleAsnGlySerLeuMetCysArgLeuGln 60
QY 181 GAAAGGTGTTATTCAGAGCTCATTCCAGCATTCAGCGGAACATAAGTGTGACATTGAT 240
DB 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATGTTAGCATCAAGCTCCCAAGATGTGCTCTCTCTTCA 300
DB 81 CysSerGlnLysValValValIlyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTCTCAGTACTCTTGGGTAACTGAGAGAAGACTTCACTCTGTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisGlu 120
QY 361 CTTCGAGGTTGGTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
DB 121 LeuAlaGlyLysIlePheAlaGlnPheSerArgCysPheProGlyLeuGlyLysSer 140
QY 421 ACTCTAGTCCCTCAGTATTCCTCAGAGCTTGTGCTTTCGCAACATGGGCAACC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTTCCCAATCTTATCTTGGCTGCGCAGAGATGCTCTCAACAAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuMet 180
QY 541 CACAGAAATGGATGGTTATGTTTAAATGCGACATATCTCTGCAAGAGCTGACTTT 600
DB 181 GlnGlnAsnLysIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTGTATTTCTGCTGCTGCTGATGATGACAGCTTTGTGGAATAATTTTG 660

DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
QY 661 CGCTGTTGGACAATATCAGTATGTTTCAATTGAGAAAGCAAAAGCTTCAATGATGTTT 720
DB 221 ProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTGAGTCACTGTTTACTGGGATCTCCGCTCGGACCATGCTATTCGCTATCATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY 781 AACAGAGTGAACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGACTTACT 840
DB 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
QY 841 ATATCTTCAACTTCAATTTTTCGGCCCAACTCTGACTATGAGAAGAAAGATTAAGAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLysIleLysAsn 300
QY 901 CAGACT 906
DB 301 GlnThr 302

RESULT 5
AAU79156
ID AAU79156 standard; protein; 665 AA.
XX
AC AAU79156;
XX
DT 02-JUL-2002 (first entry)
XX

Human dual-specificity phosphatase-3 (DSP-16) protein.

XX Human, dual-specificity phosphatase-3; DSP-16; autoimmune disease;
XX mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
XX cancer; graft-versus-host disease; allergy; metabolic disease;
XX abnormal cell growth; abnormal cell proliferation; contact inhibition;
XX cell cycle abnormality; anchorage independent cell growth; apoptosis;
XX intercellular adhesion; DSP-16 modulator; chromosome 12p.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Domain 242..251
FT /label= Active_site_domain
XX

XX MO200226997-A2.

XX 04-APR-2002.

XX 25-SEP-2001; 2001MO-US030124.

XX 26-SEP-2000; 2000US-0235487P.

XX (CEPR-) CEPTYR INC.

XX Luche RM, Wei B;

XX MPI; 2002-315802/35.

XX N-PSDB; ABR47596.

XX New DSP-16 polypeptide, useful for identifying modulators of its

XX activity, which can be used in the treatment of disorders such as

XX Duchenne muscular dystrophy, or cancer.

XX Claim 1, Fig 2; 87pp; English.

XX The present invention relates to a new polypeptide, DSP-16, having a 665
XX amino acid sequence, given in the specification, or a variant having at
XX least 50 % identical residues, which retains the ability to
XX dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
XX invention can be used for identifying agents which modulate DSP-16
XX activity, for modulation of a proliferative response in a cell, survival

CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present amino acid
 CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 CC gene located on chromosome 12p
 CC
 SQ Sequence 665 AA:

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:
Percent Similarity:	2.85e-169	665	1552.00	300
Best Local Similarity:	99.67%	1	99.67%	1
Query Match:	95.39%	0	95.39%	0
DB:	5	0	Indels:	0
			Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) x AAU79156 (1-665)

```

QY 1 ATGGCCCATGAGTGAATGGAAGTCAAAATGTTACTGAGAGTGGTGGCTCTGCTGAA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnIleValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAAAGTGGCTGTAATGATGCGGCGCATTTGGAATACATACATCC 120
DB 21 SerGlyThrGlnIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
QY 121 CACATTTTGGAAAGCCATTAATATCACTGCTCCAAAGCTTATGAAGGAGAGTGGCAACG 180
DB 41 HisIleLeuGlnAlaIleAsnIleAsnGlySerIleMetIleValArgIleGln 60
QY 181 GACAAAGTGAATTAATCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
DB 61 AspIleValIleIleThrGlnIleValIleGlnIleValIleGlnIleValIle 80
QY 241 TGGAGTGAAGAGTGGTGAATGATGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 300
DB 81 CysSerGlnIleValIleValIleValIleValIleValIleValIleValIle 100
QY 301 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAGAGAGTCAACTGTTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlnIleValIleValIleValIleValIle 120
QY 361 CTGGAGGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
DB 121 LeuAlaGlyValPheAlaGlnPheSerArgCysPheProGlyLeuGlnIleVal 140
QY 421 ACTTATGCTCTACCTGATCTTCTGAGTAACTGATCTTCTGAGTAACTGATCTTCTG 480
DB 141 ThrIleValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyPro 160
QY 481 CGAATCTTCCCAATCTTATCTTGGCTGCCAGAGATGCTTCAACAAGAGAGTGA 540
DB 161 ArgIleLeuProAsnLeuIleGlnIleGlnIleGlnIleGlnIleGlnIleGln 180
QY 541 CAGCAAGATGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
DB 181 GlnGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGln 200
QY 601 ATCCCGAGTCAATTCCTGCGTGGCTGCTGATGAGAGTGGTGGTGGTGGTGGTGGT 660
DB 201 IleProGlnIleSerHisPheLeuArgValProValAsnAspSerPheCysGlnIle 220
QY 661 CCGGTGTGGAACAATAGTAGATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 720
DB 221 ProIlePheAspIlePheSerValIlePheIleGlnIleValIleValIleValIle 240
QY 721 CTATGACACTGTTTACTCTGGATCTCCGCTCCGACATCGATCGATCGATCGATCGAT 780
DB 780

```

```

DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY 781 AAGAGAGTGAAGATGCTTTTATGATGACCTTACAGATTTTGAAGAAAGAAAGCCACT 840
DB 261 LysArgMetAspMetSerIleuAspIleValIleValIleValIleValIleValIle 280
QY 841 ATATCTCCAACTTCAATTTTCTGGCCCACTCTGAGCATGAGAGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlnIleLeuLeuAspIleGlnIleValIleVal 300
QY 901 CAGACT 906
DB 301 GlnThr 302

```

RESULT 6

AAU09946 standard; protein; 665 AA.

AAU09946;

18-JUN-2002 (first entry)

Protein sequence of human (dual specificity phosphatase) DSP-10.

Human; dual specificity phosphatase; DSP-10; cancer; epilepsy; stroke;
 neuronal degeneration syndrome; Alzheimer's disease; depression;
 schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
 osteoporosis; diabetes.

Homo sapiens.

WO200177340-A1.

18-OCT-2001.

06-APR-2001; 2001WO-EP003966.

10-APR-2000; 2000EP-00107143.

(MERE) MERCK PATENT GMBH.

Duecker K;

WPI; 2002-010917/01.

N-PSDB; AAS15768.

Novel dual specificity phosphatase polypeptides useful for treating
 cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
 disease, depression, schizophrenia, asthma and immune disorders.

Claim 2; Page 37-39; 43pp; English.

The present invention relates to a new isolated dual specificity
 phosphatase (DSP10) polypeptide, comprising a 665 residue amino acid
 sequence that is fully defined in the specification. The invention also
 provides a sequence encoded by a 3059 nucleotide sequence fully defined
 in the specification, and a sequence having at least 95 % identity to the
 polypeptide, or fragments or variants of DSP-10. The invention is useful
 for treating cancer e.g. leukemia, colon carcinoma, lung cancer,
 prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
 stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
 schizophrenia, cardiac myopathies, asthma, immune disorders,
 osteoporosis, diabetes and diabetes associated diseases. Type I diabetes,
 the invention are also useful as vaccines for inducing immunological
 response in a mammal, in disease diagnosis and in assays for screening
 agonistic or antagonistic compounds. Other uses of the invention include
 identifying membrane bound or soluble receptors, as a diagnostic reagent,
 in chromosome localisation studies, and as a valuable tool in tissue
 expression studies. The present sequence represents the dual specificity
 phosphatase, DSP-10, protein of the invention

Sequence 665 AA:

Alignment Scores:

Prod. No.:	2_85e-169	Length:	665
Score:	1552.00	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	1
Query Match:	95.39%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) x AAU09946 (1-665)

```

QY 1 ATGGCCCATGAGATGATTTGGAACCAATTTGTTACTAGAGAGTTGGTGGCTGTGGAA 60
DB 1 MetAlaHisGluMetCileGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAGTGTCTGCTAATGATGACCGGCACTTTGGTAATCAATACATCC 120
DB 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
QY 121 CACATTTTGAAGCATTATATCACTGCTCCAGCTTATGAAAGGAGTTGGCAACAG 180
DB 41 HisIleLeuGluValAlaIleAsnIleAsnCySerIleuMetIleValArgLeuGln 60
QY 181 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAACATAGTTGACATTGAT 240
DB 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleHisIleValAsp 80
QY 241 TGCAGTCAGAAAGTGTGATTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
DB 81 CySerGlnIleValIleValIleValIleValIleValIleValIleValIleVal 100
QY 301 GACTGTTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 101 AspCySerPheLeuThrValIleLeuGlyIleValGluIleValSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGAGTGGTGTGCTGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 121 LeuAlaGlyIleGlyPheAlaGluPheSerArgCySerPheProIleuGlyIleValSer 140
QY 421 ACTGTAGTCCCTACCTGCACTTCTGAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 141 ThrLeuValProThrCySerIleSerGlnProCySerLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTCTCCCAATCTTATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 161 ArgIleLeuProAsnLeuIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGly 180
QY 541 CACGAGAAATGGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 600
DB 181 GlnGlnAlaGlnIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGly 200
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 201 IleProGlnSerHisPheLeuAlaGlyValProValAsnAspPheCysGluIleLeu 220
QY 661 CGGTGTTGGAACAATCAGTATGATTTGATGAAAGCAAAAGCTCCCAATGATGATGTT 720
DB 221 ProIlePheAsnIleValIleValIleValIleValIleValIleValIleValIleVal 240
QY 721 CTAGTCACATGTTAGTGGATCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCA 780
DB 241 LeuValHisCysIleLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAla 260
QY 781 AAGAGATGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
DB 261 LysArgMetAspMetSerIleuAspGluAlaTyrArgPheValIleGlyIleGlyIleGly 280
QY 841 ATATCTCCAAATCTCAATTTTCTGAGGCACTGCTGCAATGATGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyIleLeuLeuAspIleGlyIleValIleValIleVal 300
QY 901 CAGACT 906
DB |||||

```

```

DB 301 GlnThr 302
RESULT 7
AAU05789
ID AAU05789 standard; protein; 665 AA.
XX
AC AAU05789;
XX
DT 08-MAY-2002 (first entry)
DE
XX Human protein phosphatase 7 (PP7) protein sequence.
XX
XX Human: protein phosphatase 7; PP7; immune system disorder; AIDS;
XX acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
XX Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
XX dementia; Parkinson's disease; developmental disorder; Down's syndrome;
XX cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
XX melanoma; myeloma sarcoma.
XX
XX Homo sapiens.
XX
XX Location/Qualifiers
XX Key 11..131
XX Domain /label= Rhodanese_like_domain
XX Region 15..170
XX /note= "Dual specificity protein phosphatase"
XX Region 85..298
XX /note= "VH1-type dual specificity protein phosphatase"
XX Region 158..297
XX /label= Catalytic domain DSPC
XX Domain /note= "Dual specificity phosphatase"
XX Region 220..281
XX /note= "Tyrosine specific protein phosphatases active
XX site tyr_phosphatase.prf"
XX Region 242..254
XX /note= "Tyrosine phosphatase"
XX Region 242..252
XX /note= "Tyrosine specific protein phosphatases"
XX
XX WO200210363-A2.
XX
XX 07-FEB-2002.
XX
XX 26-JUL-2001; 2001MO-US023716.
XX
XX 28-JUL-2000; 2000US-0221679P.
XX 03-AUG-2000; 2000US-022372P.
XX 10-AUG-2000; 2000US-0224309P.
XX 18-AUG-2000; 2000US-0226728P.
XX 30-AUG-2000; 2000US-0229254P.
XX 08-SEP-2000; 2000US-0231366P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;
XX Stewart EA, Gandhi AR, Patterson C, Lee EA, Hafalia AA, Lu DM;
XX Tribouley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;
XX Walla NK, Kearney L;
XX
XX MPI; 2002-188735/24.
XX N-PSDB; ABK14474.
XX
XX New protein phosphatases, useful for diagnosing, treating or preventing
XX immune system disorders (e.g. Crohn's disease), neurological disorders
XX (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
XX cancers).
XX
XX Claim 1; Page 107-108; 117p; English.
XX
XX The present invention relates to a new polypeptide, a naturally occurring
XX amino acid sequence at least 95 % identical to it, a biologically active
XX fragment of it or an immunogenic fragment of it. The polypeptides,
XX polynucleotides, agonists and antagonists are useful for diagnosing,
XX

```



```

Db      1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
QY      61 AGTGAACGGAAGAGTGTCTGTAATTGATAGCCGGCATTTGGTAATACATACATCC 120
Db      21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTrpAsnThrSer 40
QY      121 CACATTTTGGAGCCATTATATCACTGCTCCAGCTTGTGAAGGAGGATGGCAACG 180
Db      41 HisIleLeuGluAlaIleAsnIleAsnCySerIleValMetIleValArgLeuGlnGln 60
QY      181 GACAAAGTGAATTAATACAGAGCTATCCAGCATTCAGCGAAGTAAGTTGACATTGAT 240
Db      61 AspLysValLeuLeuIleThrGluLeuIleGlnHisSerIleValHisValAspIleAsp 80
QY      241 TGCAGTCAGAGGTTGATTGATACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 300
Db      81 CySerGlnIleValValValIleAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY      301 GACTGTTTCTCACTGACTTCTGGGTAACCTGGAGAAGGCTTCACTGTTCACTG 360
Db      101 AspCyPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY      361 CTTCGAGTGGTGTGCTGAGTTCCTCTGTTTCCCTGAGCTCTGTGAAGGAAATCC 420
Db      121 LeuAlaGlyGlyPheAlaGluPheSerArgCyPheProGlyLeuCySerGluGlyLysSer 140
QY      421 ACTGTAGTCCCTACCTGCATTTCTCAGCCTTGTCTTACTGTTCCCAATTTGGCCCAAC 480
Db      141 ThrLeuValProThrCyHisIleSerGlnProCyLeuProValAlaAsnIleGlyProThr 160
QY      481 CGAATCTTCCCAATCTTTACTTCTGTGCTGCACGAGATCTCTCAACAGAGCGATGA 540
Db      161 ArgIleLeuProAsnLeuValLeuGlyCySerGlnArgAspValLeuAsnLysGluLeuMet 180
QY      541 CACGAGAATGGATTGGTTATGTTAAATGCAAGCTATACCTGTCCAAAGCTGACTTT 600
Db      181 GlnGlnAsnGlyIleGlyTrpValLeuAsnAlaSerAsnThrCySerProLysProAspPhe 200
QY      601 ATCCCGAGTCTCATTTCTGCGGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Db      201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCySerGluLysIleLeu 220
QY      661 CCGTGTGTGACAAATACAGTAATTCATGAGAAAGCAAAAGCCCAATGATGATGTT 720
Db      221 ProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCyVal 240
QY      721 CTAGTCACTGTTAGTGGATCTCCGCTCCGCAACATCGCTATCGCTACATCATG 780
Db      241 LeuValHisCysLeuAlaGlyIleSerArgSerIleThrIleAlaIleAlaIleMet 260
QY      781 AAGAGATGAGACATGTTTGAATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 840
Db      261 LysArgMetSerPheMetSerLeuAspGlnAlaIleTrpArgPheValIleGluLysArgProThr 280
QY      841 ATATCTCCAAACTTCATTTTCTGGGGCAACTCTGAGCATATGAGAAAGATTAAGAAC 900
Db      281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTrpGluLysIleLysAsn 300
QY      901 CAGACT 906
Db      301 GlnThr 302

```

```

KM      Human; dual specificity phosphatase 21117; erythroid-related disorder;
KW      haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
KM      erythrocytosis; liver-related disorder; cancer.
OS      Homo sapiens.
FH      Key
FT      Domain
FT      11..131
FT      /label= Rhodanese_like_domain
FT      Region.
FT      21..24
FT      /note= "Casein kinase II phosphorylation site"
FT      91..94
FT      /note= "Casein kinase II phosphorylation site"
FT      158..297
FT      /note= "Dual specificity phosphatase catalytic domain"
FT      Region
FT      214..217
FT      /note= "Casein kinase II phosphorylation site"
FT      242..254
FT      /note= "Tyrosine specific protein phosphatase active
FT      site"
FT      Region
FT      242..254
FT      /note= "C-X5-R motif"
FT      Region
FT      266..269
FT      /note= "Casein kinase II phosphorylation site"
FT      Region
FT      369..372
FT      /note= "Casein kinase II phosphorylation site"
FT      Region
FT      421..424
FT      /note= "Casein kinase II phosphorylation site"
FT      434..437
FT      /note= "Casein kinase II phosphorylation site"
FT      Region
FT      458..461
FT      /note= "Casein kinase II phosphorylation site"
FT      Region
FT      508..511
FT      /note= "Casein kinase II phosphorylation site"
FT      Region
FT      589..592
FT      /note= "Casein kinase II phosphorylation site"
FT      Region
FT      612..615
FT      /note= "Casein kinase II phosphorylation site"
FT      Region
FT      617..620
FT      /note= "Casein kinase II phosphorylation site"
FT      Region
FT      642..645
FT      /note= "Casein kinase II phosphorylation site"
PN      US2002034807-A1.
XX
XX      21-MAR-2002.
XX
XX      23-MAR-2001; 2001US-00816494.
XX
XX      24-MAR-2000; 2000US-0191858P.
XX
PA      (MEYE/) MEYERS R A.
PI      Meyers RA;
XX
XX      MPI; 2002-351088/38.
XX      N-PDSB; ABK49402.
XX
XX      New nucleic acids, designated 38692 and 21117, encoding dual specificity
XX      phosphatases for treating cell proliferation and differentiation
XX      disorders including hematopoietic and erythroid-related disorders and
XX      cancers.
XX      Claim 8; Fig 1; 76pp; English.
XX
XX      The present invention relates to new nucleic acids designated 38692 and
XX      21117 encoding dual specificity phosphatase family members. The nucleic
XX      acid, polypeptide encoded by it, and antibody specific for the
XX      polypeptide may be used to diagnose and treat haematopoietic-related
XX      disorders such as leukaemias and autoimmune diseases, erythroid-related
XX      disorders such as anaemias and erythrocytosis, liver-related disorders,
XX      and cancers, particularly of the breast, colon, adipose, prostate and
XX      lung. The present amino acid sequence represents the human dual

```

CC specificity phosphatase 21117 protein of the invention, as described
 CC above

XX Sequence 665 AA;

Alignment Scores:

Pred. No.:	2,856-169	Length:	665
Score:	1552.00	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	1
Query Match:	95.39%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) x AAU79929 (1-665)

```

QY 1 ATGGCCCATGAGTATGTTGGAATCAATGTTACTGAGAGTGGTGGCTGCTGGAA 60
    |||||
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
QY 61 AGTGGAAAGGAAAAAGTGGCTGCTTAATGATAGCCGGCCATTTGTGGAATCAATCATCC 120
    |||||
DB 21 SerGlyThrGlnValLeuLeuIleAspSerArgProPheValGluTyAsnThrSer 40
QY 121 CACATTTTGGAAAGCATTAATCAATCTGCTCAAGCTTATGAACGGAAGTTGCAACAG 180
    |||||
DB 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgValLeuGlnGln 60
QY 181 GACCAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 240
    |||||
DB 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY 241 TGCAGTCAGAAAGGTTGAGTATGATCAATCAAGCTCCCAAGATGTTGCTCTCTGTTCA 300
    |||||
DB 81 CysSerGlnLysValValValLysAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTCTGACTGTAATCTGAGGTAATGAGAGAGGCTTCAACTCTGTTGACCTG 360
    |||||
DB 101 AspCysPheLeuThrValLeuLeuGlnLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTGAGAGTGGGTTTGGAGTCTCTGCTGTTGTTCTGCTGCTGCTGAGAGAAATCC 420
    |||||
DB 121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuGlyGlnGlyLysSer 140
QY 421 ACTCAGAGCTTCACTGATCTTCTGAGCTTGTACCTGCTTACCTGCTCAACATTGGGCCAAC 480
    |||||
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTCTCCAAATCTTATCTGGCTGCCAGCGAGATGCTTCAACAAGAGCTGATA 540
    |||||
DB 161 ArgIleLeuProAsnLeuLysLeuGlnCysGlnArgAspValIleAsnLysGlnLeuMet 180
QY 541 CACAGAAATGGGATTTGGTATGTTAAATGCAAGCTATACCTGTCGAAGCTGACTTT 600
    |||||
DB 181 GlnGlnAsnGlyIleGlyTyValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTGCGTGGCTGCTGAAATGACAGCTTTGTGAGAAATTTTG 660
    |||||
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
QY 661 CCGTGGTTGGACAAATGATGATTTATTTAGAGAAAGCAAGCCCAATGAGATGTT 720
    |||||
DB 221 ProTyrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnIleCysVal 240
QY 721 CTAGTCGACTGTAGTGGATCTCCGCTCGGCACCAATGCGTATGCGCTTAACATCATG 780
    |||||
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyIleMet 260
QY 781 AAGAGATGAGCATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTACT 840
    |||||
DB 261 LysArgMetAspMetSerLeuAspGlnAlaTyArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCAACTCAATTTCTGGGCGCAATCCCGACTGATGAGAGAAAGATTAAAGAAC 900
    |||||
  
```

DB 281 IleSerProAsnPheAsnPheLeuGlnLysLeuLeuAspTyrgLysLysIleLysAsn 300
 QY 901 CAGACT 906
 DB 301 GlnThr 302

RESULT 10
 ABB97291
 ID ABB97291 standard; protein; 665 AA.
 XX
 AC ABB97291;
 XX
 DT 27-JUN-2002 (first entry)
 XX

DE Novel human protein SEQ ID NO: 559.
 XX
 DE Human, anti-nausea; vulnerary; anti-inflammatory; immunomodulator;
 KW anti-infectivity; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag.

OS Homo sapiens.
 XX
 PN MO20022660-A2.
 XX
 XX 21-MAR-2002.
 PD
 PF 10-SEP-2001; 2001WO-US026015.
 XX
 PR 11-SEP-2000; 2000US-00659671.
 XX
 PA (HYSE-) HYSEQ INC.
 XX

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Dermanac RT,
 DR WPI; 2002-292408/33.
 DR N-PSDB; ABB97291.
 XX
 XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.
 PS Example 2; SEQ ID NO 559; 509pp; English.

CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate actin or inhibit e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 XX
 XX
 SQ Sequence 665 AA;

Alignment Scores:

Pred. No.:	2,856-169	Length:	665
Score:	1552.00	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	1
DB:	5	Indels:	0
		Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) x ABB97291 (1-665)

```

QY 1 ATGGCCCATGAGTATGTTGGAATCAATGTTACTGAGAGTGGTGGCTGCTGGAA 60
    |||||
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
QY 61 AGTGGAAAGGAAAAAGTGGCTGCTTAATGATAGCCGGCCATTTGTGGAATCAATCATCC 120
    |||||
  
```

Db	21	SerGlyThrGluIuValIleuLeuIleAspSerArgProPheValGluTyraAnthSer	40	XX	Key	Location/Qualifiers
QY	121	CAATTTTGGAGCCATTATATCACTGCTCCAGGTTATGAAGCAAGTTGCAACAG	180	FT	Modified-site	12 /note= "O-phosphorylated"
Db	41	HisIleLeuGluIuValIleAenIleAsnCySerIuLeuMetLysArgIleuGln	60	FT	Modified-site	21 /note= "O-phosphorylated"
QY	181	GACAAAGTTAAATTCAGAGCTCATCCAGCATTCAGCAAAATTAAGTTGACATTGAT	240	FT	Modified-site	23 /note= "O-phosphorylated"
Db	61	AspIuValIleuIleThrGluLeuIleGlnHisSerAlaIuValAspIleAsp	80	FT	Modified-site	38 /note= "N-glycosylated"
QY	241	TGCAGTCAGAGGTTGATTATGCATCAAGCTCCCAAGATGTGCTCTCTCTCA	300	FT	Modified-site	38 /note= "N-glycosylated"
Db	81	CySerGlnIuValValValIuTyraGlnSerSerGlnAspValAlaSerIuSerSer	100	FT	Modified-site	49 /note= "N-glycosylated"
QY	301	GACTGTTTTCCTGCTGATCTTCTGGGTTAACTGAGAAAGCTTCAACTTGTTCACCTG	360	FT	Modified-site	72 /note= "O-phosphorylated"
Db	101	AspCyPheLeuThrValIleuLeuGluIuValLeuGluIuLysSerPheAsnSerValHisIleu	120	FT	Modified-site	82 /note= "O-phosphorylated"
QY	361	CTTGCAAGTGGGTTTGCTGAGTCTCTCGTTGTTTCCCTGCTGCAAGGAAATCC	420	FT	Region	85..298 /note= "O-phosphorylated"
Db	121	LeuAlaGluIuValPheAlaGluPheSerArgCyPheProIuLeuCyGluIuLysSer	140	FT	Modified-site	91 /note= "VHL-type dual specificity phosphatase signature"
QY	421	ACTGTAGTCCCTACCTGATTTTCAAGCCTTGCTTACTGTTGCCAATGGGCAACC	480	FT	Modified-site	190 /note= "O-phosphorylated"
Db	141	ThrIuValProThrCySerIuSerGlnProCyLeuProValAlaAsnIleGlyProThr	160	FT	Modified-site	190 /note= "N-glycosylated"
QY	481	CGAATCTTCCCAATCTTTATCTTGGCTGCGACGAGATGCTCTCAACAAGAGCTGATA	540	FT	Modified-site	212 /note= "N-glycosylated"
Db	161	ArgIleLeuProAsnLeuTyrluGluIuCyGlnArgAspValIleuAsnIuLysIuLeuMet	180	FT	Modified-site	214 /note= "N-glycosylated"
QY	541	CAGCAGATGGGATTGTTATGTTAAATGCGAGCTATCTGTCCAAAGCCTGACTTT	600	FT	Active-site	220..280 /note= "O-phosphorylated"
Db	181	GlnGlnAsnGluIleGlyTyrluValIleuAsnAlaSerAnthCySerPolySproAspPhe	200	FT	Region	237..278 /note= "Tyrosine specific protein phosphatase"
QY	601	ATCCCGAGTCTGATTTCTGCGTGCCTGTGATGATGACGCTTTGTGAAATTTTG	660	FT	Modified-site	256 /note= "xy phosphatase signature"
Db	201	IleProGluSerHisPheLeuArgValProValAsnAspSerPheCyGluIuLysIleu	220	FT	Modified-site	280 /note= "O-phosphorylated"
QY	661	CGGTGGTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTTCAATGGATGTT	720	FT	Modified-site	300 /note= "O-phosphorylated"
Db	221	ProThrLeuAspLysSerValAspPheIleGluIuValAlaLysAlaSerAsnGlyCyVal	240	FT	Modified-site	369 /note= "N-glycosylated"
QY	721	CTAGTCACGTTAGCTGGGATCTCCGCTCCGCCACCACTGCTTGGCTCATCATCAG	780	FT	Modified-site	393 /note= "O-phosphorylated"
Db	241	LeuValHisCySerLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrluMet	260	FT	Modified-site	421 /note= "O-phosphorylated"
QY	781	AAGAGATGAGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGACTACT	840	FT	Modified-site	422 /note= "O-phosphorylated"
Db	261	LysArgMetAspMetSerIleuAspGluAlaTyrluArgPheValLysGluLysArgProThr	280	FT	Modified-site	434 /note= "O-phosphorylated"
QY	841	ATATCTTCCAAACTTCAATTTTCTGGGCAACTCTTGAGCTATGAGAAAGAAATTAAACA	900	FT	Modified-site	439 /note= "O-phosphorylated"
Db	281	IleSerProAsnPheAsnPheLeuGluIuLeuLeuAspTyrluLysLysIleLysAsn	300	FT	Modified-site	468 /note= "O-phosphorylated"
QY	901	CAGACT 906		FT	Modified-site	471 /note= "O-phosphorylated"
Db	301	GlnThr 302		FT	Modified-site	479 /note= "O-phosphorylated"
RESULT 11				FT	Modified-site	528 /note= "O-phosphorylated"
AAB20325				FT	Modified-site	590 /note= "O-phosphorylated"
AAB20325 standard; protein; 666 AA.				FT	Modified-site	590 /note= "O-phosphorylated"
XX				FT	Modified-site	597 /note= "O-phosphorylated"
XX				FT	Modified-site	605 /note= "O-phosphorylated"
XX				FT	Modified-site	610 /note= "O-phosphorylated"
DT	29-MAY-2001 (first entry)			FT	Modified-site	613 /note= "O-phosphorylated"
XX				FT	Modified-site	618 /note= "O-phosphorylated"
XX				FT	Modified-site	628 /note= "O-phosphorylated"
DE	Human protein phosphatase and kinase protein-4.			FT	Modified-site	641 /note= "O-phosphorylated"
XX				FT	Modified-site	
KW	Protein phosphatase and kinase protein; PPK-4; human;			FT	Modified-site	
KW	gastric intestinal disorder; immune system disorder; neurological disorder;			FT	Modified-site	
XX	cell proliferative disorder; cancer; diagnosis; therapy.			FT	Modified-site	
OS	Homo sapiens.			FT	Modified-site	

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y;
 XX MPI; 2003-395539/38.
 DR N-PSDB; ADA53105.
 XX
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 PS Claim 14; SEQ ID NO 2312; 205bp; English.
 XX
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.
 XX
 SQ Sequence 665 AA;
 Alignment Scores:
 Pred. No.: 1.83e-168 Length: 665
 Score: 1545.00 Matches: 299
 Percent Similarity: 99.34% Conservative: 1
 Best Local Similarity: 99.01% Mismatches: 2
 Query Match: 94.96% Indels: 0
 DB: Gaps: 0
 US-10-029-345A-108_COPY_538_1443 (1-906) X ADA54744 (1-665)
 QY 1 ATGCGCCATGAGATGATGGAATCAATTTGTACTGAGAGGTTGGCTGCTGGA 60
 DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGln 20
 QY 61 AGTGAACGGGAAAGAGTGTGCTTATTTGATAGCGGCGCATTTTGTGAATACATAC 120
 DB 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGlnTyrAsnTrSer 40
 QY 121 CACATTTTGGAAAGCATTAATATCACTGCTCCAGCTTATGAAAGGAGTTGCAAC 180
 DB 41 HisIleLeuGlnAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 60
 QY 181 GACAAAGTGTATTTACAGAGCTCATCCAGCATTCAGCGGAACATTAAGTTGAC 240
 DB 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisLysValAsp 80
 QY 241 TGCAGTCAGAGGTTAGTTAGATCAAGCTCCAGAGTGTGCTGCTCTCTTCA 300
 DB 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeu 100
 QY 301 GACTGTTTCTCACTGACTTCTGGGTAAACCTGGAAGAGCTTCACTCTGTTCA 360
 DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
 QY 361 CTTCGAGGTGGGTTGCTGAGTTCTCTGTTTCTGCTGGCTCTGTGAAGAAATCC 420
 DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyLysSer 140
 QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTGCTTACCTGTTGCGCAACATTGG 480
 DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProAlaAlaAsnIleGlyProThr 160
 QY 481 CGAATTTCTCCCAATCTTTATCTTTGCTGCGCAGCGAGATGCTCTCAACAAGAG 540
 DB 161 LeuIleLeuProAsnLeuThrLeuGlyCysGlnAspValLeuAsnLysGluLeuMet 180
 QY 541 CAGCAGAAAGGAGTTGGTATGTGTATTAATGCCAGCTATACCTGTCAAGAGCTTT 600
 DB 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
 QY 601 ATCCCGAGTGTATTTCTGCTGCTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG 660

DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
 QY 661 CCGTGGTGGACAAATCAGTATGATTTTCATTGAGAAAGCAAAAGCTCCCAATGATGT 720
 DB 221 ProThrLeuAspLysSerValAspPheIleGlyLysValAlaLysAsnGlyCysVal 240
 QY 721 CTATGCACTGTTTGGCTGGATCTCCGCTCCGACCATGCTTGTGCAAGAAAGACCTACT 780
 DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 QY 781 AAGAGGATGACATGCTTTAGATGAGCTTACAGATTTGTGAAAGAAAGACCTACT 840
 DB 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGluLysArgProThr 280
 QY 841 ATATCTCCAAATCTCAATTTTCTGGCCCACTCTGAGATTAAGAAAGATTAAGAAC 900
 DB 281 IleSerProAsnPheAsnPheLeuGlnLeuLeuAspTyrGluLysLysIleLysAsn 300
 QY 901 CAGACT 906
 DB 301 GlnThr 302
 RESULT 13
 AAU79161
 ID AAU79161 standard; protein; 665 AA.
 XX AAU79161;
 AC
 XX
 XX 02-JUL-2002 (first entry)
 DT
 XX
 DE Human dual-specificity phosphatase-3 (DSP-16) mutant protein #1.
 XX
 XX Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KW cancer; graft-versus-host disease; allergy; metabolic disease;
 KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KW intercellular adhesion; DSP-16 modulator; mutant; mutain.
 XX
 OS Homo sapiens.
 XX Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 213 /note= "wild-type Asp substituted by Ala"
 FT
 XX
 PN WO200226997-A2.
 XX
 PD 04-APR-2002.
 XX
 PF 25-SEP-2001; 2001WO-US030124.
 XX
 PR 26-SEP-2000; 2000US-0235487P.
 XX
 PA (CEPT-) CEPTYR INC.
 XX
 PI Luche RM, Wei B;
 XX
 XX MPI; 2002-315802/35.
 DR
 PT New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 PS Claim 46; Page; 87pp; English.
 XX
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated, mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity, for modulation of a proliferative response in a cell, survival

CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present amino acid
 CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 CC mutant protein #1. Note: This sequence is not shown in the specification
 CC but is derived from the wild-type human DSP-16 (AAU79156) protein given
 CC in figure 2 of the specification
 XX
 SQ Sequence 665 AA;

Alignment Scores:

Pred. No.:	2,38e-168	Length:	665
Score:	1544.00	Matches:	299
Percent Similarity:	99.34%	Conservative:	1
Best Local Similarity:	99.01%	Mismatches:	2
Query Match:	94.90%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-108_COPY_538_1443 (1-906) x AAU79161 (1-665)

```

QY 1 ATGGCCCATGAGATGATGGAAGTCAATGTTTCTGAGAGGTGGCTGCTGCGAA 60
Db 1 MetAlaHLeuMetLeuGlyThrGlnLeuValThrGlnLeuValAlaLeuLeuGln 20
QY 61 AGTGAACGGAAGAAAGTCTGCTAATGATGCGCGCATTTGTGAATCAATACATCC 120
Db 21 SerGlyThrGlnLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
QY 121 CACATTTTGAAGCATTATATCACTGCTCCAGCTTATGAGAGGAGTTGCAACG 180
Db 41 HisLeuLeuGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
QY 181 GACAAAGTGTAAATTAACAGAGCTTATGAGAGGAGTTGCAACATGATGATGAT 240
Db 61 AspLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
QY 241 TGCAAGTGAAGAGTGTGATGTTTACGATCAAGCTCCAGAGTGTGCTCTCTTCA 300
Db 81 CysSerLeuValValValValValValValValValValValValValValValVal 100
QY 301 GACTGTTTCTCACTGATCTTGGGTAACCTGGAAGAGCTTCACTGTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
QY 361 CTTCGAGGTGGGTTGCTGAGTTCTCTGTTGTTCCCTGAGCTCTGAGAGGAAATCC 420
Db 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyLeuSer 140
QY 421 ACTGATGCTTACCTGAGATTTCTCAGGCTTCTTACTGTTGCAACATGAGCCAC 480
Db 141 ThrLeuValProThrCysLeuSerGlnProCysLeuProValAlaSerLeuGlyProThr 160
QY 481 GGAATTTCTCCCAATCTTTATCTTGGCTGCCAGAGATGCTCCCAAGAGAGGAGTGA 540
Db 161 ArgGlyLeuProAsnLeuValLeuGlyCysGlnArgPheValLeuAsnGlyLeuLeu 180
QY 541 CAGCAGAAATGGGATTTGTTATGTTTAAATGCAAGCTTACTGTTCAAGGCTGACTTT 600
Db 181 GlnGlnAsnGlyLeuGlyValLeuLeuAsnAlaSerAsnThrCysProLeuProAspPhe 200
QY 601 ATCCCGAGGCTCAATTTCTGCGTGCCTGATGATGACAGCTTTTGGAGAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAlaSerPheCysGlnLeuLeu 220
QY 661 CCGTGTGTGACAATAGTAGATTTCAATGAGAAAGCAAGGCTCCCAATGATGTTGTT 720
Db 221 ProThrPheAspLeuValAspPheLeuGlyValValValValValValValValVal 240
QY 721 CTAGTGACGTGTTAGCTGGAGATCTCCCGCTCCGACCATCGATCGATGCTTACATCAG 780

```

```

Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
QY 781 AAGAGATGAGACATGCTTTTATGATGACCTTACGATTTTGTGAAGAAAAAGACCTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValIysGlnLysArgProThr 280
QY 841 ATATCTCCAACTTCAATTTTCTGGCCCAACTCCTGAGCATATGAGAAAGATTAAGAC 900
Db 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysIleLysAsn 300
QY 901 CAGACT 906
Db 301 GlnThr 302

```

RESULT 14
 ABR52352
 ABR52352 standard; protein; 665 AA.
 ABR52352;
 19-JUN-2003 (first entry)

Protein relating to the invention SEQ ID NO: 42.
 antiproliferative; hepatotropic; nephrotropic; antiarthritic;
 antiproliferative; cardiac; cytosolic; gene therapy; liver disease;
 proliferative disorder; renal failure; cardiovascular disorder;
 immunological disorder; arthritis; psoriasis; congenital heart defect;
 congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
 Homo sapiens.
 MO200257460-A2.
 25-JUL-2002.
 20-DEC-2001; 2001WO-US050459.
 20-DEC-2000; 2000US-0256868P.
 30-MAR-2001; 2001US-0280186P.
 01-MAY-2001; 2001US-0287735P.
 05-JUN-2001; 2001US-0295848P.
 25-JUN-2001; 2001US-0300465P.
 (BRIM) BRISTOL-MYERS SQUIBB CO.
 Jackson DG, Feder J, Nelson T, Mintier G, Ramamathan C, Lee L,
 Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
 Krystek S, Mcatee P, Suchard S, Banas D;
 WPI; 2002-599721/64.
 N-PSDB; ACC60521.
 Novel polynucleotides encoding human phosphatase polypeptides useful in
 the prevention or treatment of e.g. proliferative and cardiovascular
 disorders.
 Claim 5; Fig 5; 801pp, English.
 The invention relates to a novel isolated nucleic acid comprising a
 polynucleotide having a nucleotide sequence selected from 40
 polynucleotides fully defined in the specification. The polynucleotide of
 antiarthritic, antiproliferative, hepatotropic, nephrotropic,
 CC antiproliferative, cardiac, and cytosolic activity. The
 CC polypeptide may have a use in gene therapy. A polynucleotide or
 CC ameliorating a medical condition, e.g. a proliferative disorder, treating or
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present

CC sequence is used in the exemplification of the invention
XX
SQ Sequence 665 AA:

Alignment Scores:

Pred. No.:	Length:	665
Score:	1544.00	299
Percent Similarity:	99.34%	1
Best Local Similarity:	99.01%	2
Query Match:	94.90%	0
DB:	5	0

US-10-029-345A-108_COPY_538_1443 (1-906) x ABR52352 (1-665)

```

QY 1 ATGGCCCATGAGATGTTGGAATCTCAATGTGTACTGAGAGGTTGGTGGCTTCTGCTGAA 60
DB 1 MetAlaH16GluMet1IleGlyThrGln1IleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAAAGTGTCTAATTGATGAGCCGCGCATTTGTGGAATACATACATCC 120
DB 21 SerGlyThrGluValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
QY 121 CACATTTGGAGACCATTAATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 180
DB 41 HisIleLeuGlnAlaIleAen1IleAenCysSerLeuMetLeuValArgArgLeuGln 60
QY 181 GACAAAGTGTATTTACAGAGCTCATCCAGCATTCAGCCAAACATTAAGTTGACATTGAT 240
DB 61 AspLyValLeuLeuLeuThrGluLeuLeuIleGlnHisSerAlaLeuHisValAspIleAsp 80
QY 241 TGACAGCAGAGAGTGTAGTTTACGATCAAGCTCCCAAGATGTTGGCTCTCTCTTCA 300
DB 81 CysSerGlnValValValValValValValValValValValValValValValValVal 100
QY 301 GACTGTTTCTCAGTCTACTTCTGGGTAAACTGAGAGAGAGCTTCACTCTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTGGCAGGTTGGTTTGTGAGTTCTCTGTTTCTCTGTTTCTCTGTTGAGAGAAATGC 420
DB 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProLysGluGlyLysSer 140
QY 421 ACTCTAGCTCCCTACCTGATTTCTCAGACCTTGCTTACCTGTTGCCAAACATTTGGGCCAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTTCTTCCCAATTTTATCTTGGCTGCGCAGCAGAGATGCTCTCAACAAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuThrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuMet 180
QY 541 CAGCAGATGGGATTTGGTTATGTGTAAATGCCAGTATACCTGTTCCAAAGCTGACTTT 600
DB 181 GlnGlnAlaAsnGlyIleGlyThrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTGATTTCCGGGTGCTGCTGGAATGACAGCATTGTTGTGAGAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuAlaGlyValProValAsnAspSerPheCysGluLysIleLeu 220
QY 661 CCGTGGTTGACAAATCAGTAGATTTCATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 221 ProThrLeuAspLysSerValAspPheIleGluLysValAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTAGTGCATCTTTAGCTGGGATCTCCGCTCCGCCCACTGCTTATCCGCTTACATCANG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIleMet 260
QY 781 AAGAGAGTGAACATGCTCTTAGAGAGAGCTTACAGATTGTGTAAGAGAGAGAGAGAGAG 840
DB 261 LysArgMetAspMetSerLeuAspGluAlaThrArgPheValLysGluLysArgProThr 280
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCCACTCTGAGCATAGAGAGAGAGAGAGAGAGAG 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAlaThrGluLysLysIleLysAsn 300

```

```

QY 901 CAGACT 906
DB 301 GlnThr 302

RESULT 15
ID AAM25744
AAM25744 standard; protein; 672 AA.
AC AAM25744;
XX 16-OCT-2001 (first entry)
DT XX
DE Human protein sequence SEQ ID NO:1259.

KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW anti-inflammatory; antirheumatic; antidiabetic; immunosuppressive;
KW antibacterial; endocrine; cardiac; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiparasitic; haemostatic; vulnery; antilicer; osteopathic; eczema;
KW dermatological; antiallergic; antiseptic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.

XX Homo sapiens.
OS
XX WO200153455-A2.
PN
XX 26-JUL-2001.
PD
XX 22-DEC-2000; 2000WO-US035017.
PF
XX 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
XX
XX (HYSR-) HYSER INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI WPI; 2001-457603/49.
XX DR N-PSDB; AAM99685.
XX
XX Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
PT
XX Claim 20; Page 260; 1217pp; English.
PS
XX AAM99166 to AAM99904 encode the human proteins given in AAM25225 to
CC AAM25965. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: anti-inflammatory; antirheumatic;
CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiac;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiparasitic; haemostatic; vulnery;
CC antilicer; osteopathic; dermatological; antiallergic; antiseptic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic

```

CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
XX
SQ Sequence 672 AA;

Alignment Scores:

Pred. No.:	2,39e-168	Length:	672
Score:	1544.00	Matches:	299
Percent Similarity:	99.34%	Conservative:	1
Best Local Similarity:	99.01%	Mismatches:	2
Query Match:	94.90%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345a-108_copy_538_1443 (1-906) x AAM25744 (1-672)

QY 1 ATGGCCCATGAGATGATGGAATCGAAATGTTACTGAGAGGTGTGGCTCTGCTGAA 60
DB 8 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgGlyValAlaLeuLeuGlu 27
QY 61 AGTGGAAACGGAAGAAAGTGTGCTAAATTGATAGCCGCGCAATTGTGGAATTAATAATACATCC 120
DB 28 SerGlyThrGlnLysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 47
QY 121 CACATTTTGAAGACCATTAATATCACTGCTCCAGCTTGAAGCGAAGGTGCAACAG 180
DB 48 HisIleLeuGluAlaIleAsnIleAsnGlySerLysLeuMetLysArgLeuGlnGln 67
QY 181 GACAAAGTGAATTAATCAAGCTCATCAGCATTCAGCGAAACATTAAGGTGACATTGAT 240
DB 68 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 87
QY 241 TGCAGTCAGAGGTGTAGATTAGATCAAGCTCCCAAGATTGCTCTCTCTTCA 300
DB 88 CysSerGlnLysValValIleValIleValIleValIleValIleValIleValIleValIleVal 107
QY 301 GACTGTTTCTCACTGACTCTTGGGTAACCTGAGAGAGCTTCAACTGTGCTGCTTCA 360
DB 108 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysPheAsnSerValHisIleu 127
QY 361 CTTCAGAGTGGGTTTGGCTGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
DB 128 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlyLysSer 147
QY 421 ACTGTGCTCCCTACCTGCAATTTCTCAGCTTGTACTGCTGCTCAACATGGGCGAAC 480
DB 148 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 167
QY 481 CGAATTCCTCCCAATCTTAATCTTGGAGCAAGAGATGTCCTCAACAAGAGCTGATA 540
DB 168 ArgIleLeuProAsnLeuLysLeuGlyCysGlnArgAspValLeuAsnLysGluLeuMet 187
QY 541 CAGCAAGATGGATGGTTATGTGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 600
DB 188 GlnGlnAsnGlyIleGlyLysValLeuAsnAlaSerAsnThrCysProLysProAspPhe 207
QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTAATGACAGCTTTGTGAGAAATTTTG 660
DB 208 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 227
QY 661 CCGTGGTTGCAAAATGATGATTTATTTAGAGAAAGCAAAAGCTCCAAATGATGTT 720
DB 228 ProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 247
QY 721 CTAGTGACTGTGTAGCTGGATCTCCCGCTCCGCAACATGGCTATGCTTACATCATG 780
DB 248 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaLysIleMet 267
QY 781 AAGAGATGACATGCTCTTATGATGAAGCTTACAGATTGTGAAAGAAAGAAAGCTTACT 840
DB 268 LysArgMetCysPheSerLeuAspGluAlaLysArgPheValLysGluLysArgProThr 287
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCGAACTCTGACTATGAGAGAGATTGAAGAC 900

DB	288	IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysLysIleLysAsn 307
QY	901	CAGACT 906
DB	308	GlnThr 309

Search completed: June 21, 2004, 12:34:22
Job time : 51.4105 secs

us-10-029-345a-108_copy_538_1443.rabb

Sequence 2, Appl1
Sequence 2, Appl1
Sequence 680, App

Sequence 26. Appl

Sequence 2, Appl1
Sequence 240, App
Sequence 247, App

Sequence 703, APP
E4304 M
Costo...

Sequence 681, APF
Sequence 231? Ar

Sequence 1259, App. Sequence 682, App.

Sequence 256, App

Sequence 700, App

Sequence 7, Appli

Sequence 15, Appl

sequence 11, Appl

Sequence 14, Appl

Sequence 701, APF

Sequence 6, Appli

Sequence 47, Appl
Commence 120 Apr

Sequence 2, Appl. Sequence 164, Ann

sequence 2, applied
sequence 25. Appl

sequence 805, App

sequence 805, App

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 258
LENGTH: 662
TYPE: PRN
ORGANISM: Homo sapiens
US-10-072-012-258

Alignment Scores:

Pred. No.: 6.02e-152 Length: 662
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 0
Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-072-012-258 (1-662)

QY 1 ATGGCCCATGATGATTTGAACTCAATTTGTTACTGAGAGGTTGGCTTGTGGA 60
DB 1 MetAlahisglumetilegilythrghinilevalthrghinleuvalAlaleuLeuGlu 20
QY 61 AGTGAACGGAAGAGTGTCTATTTGATGCGGCGCATTTGGAATCAATACATCC 120
DB 21 SerGlyThrhGluValleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
QY 121 CACATTTTGAAGCATTAATATATCACTGCTCAAGCTTATGGAAGGAGGTTGCAACG 180
DB 41 HislleuLeuGluAlaleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
QY 181 GACAAAGTGTAAATTAAGAGCTATCCAGCATTCAGCGAAACATTAAGTTGACATGAT 240
DB 61 AspIysValleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
QY 241 TGAGTCGAAGAGTTGATTTAGATCAAGCTCCCAAGATGTTGCTCTCTTCA 300
DB 81 CysSerGlnIysValIvalIyraSpGlnSerSerGlnSpValAlaserLeuSerSer 100
QY 301 GACTGTTTCTCACTGATCTTGGGTAATCGAGAGAGCTTCACTGTTCACTG 360
DB 101 AspPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
QY 361 CTTCGAGGTGGTTTCTAGATTCTCTGTTTCCCTGAGCTCTGGAAGAAATCC 420

DB 121 LeuAlaGlyIyIyPheAlaGluPheSerArgCysPheProGlyLeuCyGluGlyIySer 140
QY 421 ACTTATGCTCTTACCTGCAATTTCTGAGCTTCTTACCTGTTGGCAATTTGGCCCAAC 480
DB 141 ThrLeuValProThrCyIleSerGlnProCysLeuProValAlaleuLeuLeuLeu 160
QY 481 CGAATTTCCCAATCTTATTTATTTGCTGCGCCAGAGATGCTTCAACAAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
QY 541 CAGCAAGATGGGATTTGTTATGTTAAATGCGATATACCTGTCCAAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyIyValleuAsnAlaserAsnThrCysProIysProAspPhe 200
QY 601 ATCCCGAGTCTCAATTTCTGCGGTCCTGTAATGACAGCTTTTGTGAGAAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIyIleLeu 220
QY 661 CCGTGTGGCAATCATGATGATTTCAATGAGAAAGAAAGCCCAATGAGATGTT 720
DB 221 ProIleuAsnIySserValAspPheIleGlyIySalIyAlaserAsnIyCysVal 240
QY 721 CTAGTCACCTGTTAGCTGAGATCTCCGCTCCGACCATGCTATGCTATGCTTACATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIyIleMet 260
QY 781 AAGAGATGAGACATGCTTTTATGATGAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
DB 261 LysArgMetAspSerSerLeuAsnSpGlnAlaIyArgPheValIyGluIyAsnArgProthr 280
QY 841 ATATCTCAAACTTCAATTTTCTGAGGCACTCCGATGATGATGAGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyIleuLeuAsnIyIyGluIySylIyIyAsn 920
QY 901 CAGACT 906
DB 301 GlnThr 302

RESULT 2

US-09-816-494-2
Sequence 2, Application US/09816494
Patent No. US20020034807A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,858
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for windows Version 4.0
SEQ ID NO 2
LENGTH: 665
TYPE: PRN
ORGANISM: Homo sapiens
US-09-816-494-2

Alignment Scores:

Pred. No.: 6.03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 0
Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x US-09-816-494-2 (1-665)

QY 1 ATGGCCCATGATGATTTGAACTCAATTTGTTACTGAGAGGTTGGCTTGTGGA 60
DB 1 MetAlahisglumetilegilythrghinilevalthrghinleuvalAlaleuLeuGlu 20

Oy		61	ACTGGAACGGAAAAAGTGTCTGTAATGATAGCGGCCCATTTGTGGAAATACATACTC	120
Db		21	SerglyThnGlulysValLeuLeuIleAepSerArzProPhaValGluTyxAnthrSer	40
Oy		121	CACATTTCGGAAGCCATTATATATCACTGCCTCCAACTTATAGAAGGAGGTTCGAACAG	180
Db		41	HstIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgHgleuGln	60
Oy		181	GACAAAGTGTTAATTAACAGAGCTCATCCAGCATTCAGCGAAAATATAGTTGACATTGAT	240
Db		61	AapLySvalIleuIleThrcIuleuIleGlnHisSerAlbShibalyValAapIleAap	80
Oy		241	TGCAGTCAGAAGCTGTAGTTATACGATCAAAGCTCCCAAGATGTTGCTCTCTCTTCA	300
Db		81	CysSerGlnLysValValValTyxApGInSerSerGlnAapValAlaSerLeuSerSer	100
Oy		301	GACGTGTTTTCTCACGTACTTCTCGGTTAACTCGAABAAGCTTCAACTCTGTTCACCTG	360
Db		101	AapCySpHeLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHIsleu	120
Oy		361	CTTGCAGGAGGTGTTGCTAGTTCTCTCGTGTGTTTCCCTGCGCTCTGTGAAGAAATCC	420
Db		121	LeuAlaGlyGlyPheAlaIeAsnLeuPheStrArgCySpHeProGlyLeuCySeGluGlyLysSer	140
Oy		421	ACTGTAGTCCCTAACCTGCATTTCTCAGCCTGTGTTACTGTGTCCAAATTCGGCCACC	480
Db		141	ThrLeuValProThrCysIleSerGlnProCySleuProValAlaAsnIleGlyProThr	160
Oy		481	CGAATTTCTCCCAATCTTTATCTTGCTGCTCCAGCGAATGTCTCAACAAGAGCTGATA	540
Db		161	ArgIleLeuProAsnLeuTyxLeuGlyCySGlnArgAPValLeuAsnLysGluLeuMet	180
Oy		541	CAGCAGAAATGGGATTGGTATGNGTTAAATGCGCAGCATATACCTGTCCAAGCTGACTTT	600
Db		181	GlnGlnAsnGlyIleGlyTyxValLeuAsnAlaSerAsnthrCysProLysProAapPhe	200
Oy		601	ATCCCCAGAGTCTCATTTCTCGCGGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG	660
Db		201	IleProGlnSerHisPheLeuArgValLProValAlaAspSerPheCysGlnLysIleLeu	220
Oy		661	CCGTGGTTGGACAAATCAGTAGATTTTCAATTGAGAAAAGCAAGCTTCCATGAGATGTGT	720
Db		221	ProtPLeuaAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCySVal	240
Oy		721	CTAGTGCACACTGTTTGGCTGGAGATCTCCGCGCTCCGACCATGGCTATGCGCTACATCATG	780
Db		241	LeuValHisCySLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyxIleMet	260
Oy		781	AAGAGGATGACATGTCCTTAGATGAGCTTACAGATTTGTGAAAAGAAAAGACCTTACT	840
Db		261	LysArgMetAapMetSerLeuAspGlnAlaTyxArgPheValLysGluLysArgProThr	280
Oy		841	ATATCTCCAAACTTCATTTTCTCGGGCCAACTCTTGAGCTATGAGAAAGATTAAAGAC	900
Db		281	IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyxGluLysIleLysAsn	300
Oy		901	CAGACT 906	
Db		301	GlnTr 302	
 RESULT 3 US-09-964-277-2 ; Sequence 2, Application US/09964277 ; Patent No. US20020137170A1 ; GENERAL INFORMATION: ; APPLICANT: Luche, Ralf M. ; APPLICANT: Wei, Bo ; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE ; FILE REFERENCE: 200125_434 ; CURRENT APPLICATION NUMBER: US/09/964,277 ; CURRENT FILING DATE: 2001-09-25 ; NUMBER OF SEQ ID NOS: 22 ; SOFTWARE: FastSeq for Windows Version 4.0				

/	SEQ ID NO 2
/	LENGTH: 665
/	TYPE: PRT
/	ORGANISM: Homo sapiens
US-09-964-277-2	
Alignment Scores:	
Pred. No.:	6,03e-152
Score:	1552.00
Percent Similarity:	99.67%
Best Local Similarity:	99.34%
Query Match:	95.39%
DB:	9
	Gaps: 0
US-10-029-345A-108_COBY_538_1443 (1-906) x US-09-964-277-2 (1-665)	
QY	1 ATGGCCCATGATGATTTAGAACTCAATTGTACTGAGAGGTGGTGCCTGCTGA 60
Dd	1 MetalaHISGLMETILEGIYThrIntIleValIThrGluArgLeuValAlaLeuLeuGlu 20
QY	61 AGTGGAACGAAAAAGTCCTGCTTAATTGATACCGGCCCTTTGTGGATAACAATACATC 120
Dd	21 SerGIythrGluValLeuLeuLeuIleAspSerAlaArgProPheValIGlutrykAntHiset 40
QY	121 CACATTTTGAAGCCATTATATATCAACTGCTCCAAGCTTATAGAAGCAAAGTTGCAACG 180
Dd	41 HisIleLeuGluAlaIleAlaAsnIleAsnCysSerLysLeuMetLysArgTrgLeuGlnGln 60
QY	181 GACAAAGTGTAAATTACACAGCTCATCCAGCATTTCAAGCAAAACATPAAGTTGATTTGAT 240
Dd	61 AspLysValIleuIleThrGluLeuIleGlnHisSerAlaLysIleValAspIleAsp 80
QY	241 TGCAGTCAAGAGGTGTGATTTACAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTGA 300
Dd	81 CysSerGlnLysValValValValTykAspGlnSerSerGlnAspValAlaSerLeuSer 100
QY	301 GACTGTTTTCTCACGTGACTTCTGGGGTAACTGGAAGAGCTTCAACTCTGTTCACCTG 360
Dd	101 AspCysPheLeuThrValLeuLeuGlyLysValLeuGluLysSerPheAsnSerValHisLeu 120
QY	361 CTGGAGAGTGGGTGTGCTGAGTCTCTCGTGTCTTCCCTGGGCTCTGTGAAGAAAATCC 420
Dd	121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
QY	421 ACTTAGTCCCCTACCTGCATTTCTCAGCCTTCTTACCTGATGCTTACCCTGCAAAAGCTGACTT 480
Dd	141 ThrIeuValProtnHcyAlIeserGlnProCysIeuProValAlaAsnIleGlyProtn 160
QY	481 CGAATCTTCCCAATCTTATCTTGGCTGCCAGCGAGATGTCTTCAACAGAGCTGATA 540
Dd	161 ArgIleLeuProAsnLeuTykIeuGlyCysGlnIleGAspValIleuAsnLysGluLeuMet 180
QY	541 CAGCAGAATGGGATGTGTTATGTGTTAAATGCCACGCTATACCTGTCCAAAAGCTGACTT 600
Dd	181 GlnGlnAsnGlyIleGlyTykValIleuAsnAlaSerAsnThrCysPheProLysProAspPhe 200
QY	601 ATCCCCGAGCTCATTTCTCGTGGTGGCTGTGAATGACAGCTTTTGTGAGAAAAATTTTG 660
Dd	201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
QY	661 CCGTGGTTGCAAAATCAGTATGATTTTCAATTGAGAAAGCAAAAGCCTTCAATGGATGTGT 720
Dd	221 ProtyrIeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
QY	721 CTAGTGCACTGTTTAGCTGGGATCTCCGGCTCCGCCACATGGCTATGCCCTTACATCAG 780
Dd	241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaIleAlaIleAlaTykIleMet 260
QY	781 AAGGGAAGGACATGCTTTTGAATGATACCTTACATATTTGTGAAGAAAAAGCACTACT 840
Dd	261 LysArgIeuAspMetSerIeuAspIleAlaTykArgPheValLysGluLysArgProtn 280
QY	841 ATATCTCAAACTTCAATTTTCTGGGCCAACCTCTGAGCTATAGAGAAAGATTAAGAAC 900

Db 281 |||||
QY 901 CAGACT 906 |||||
Db 301 GlnThr 302

RESULT 4

US-10-072-012-680
; Sequence 680, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernov, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patlurajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Futak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Rieger, Denise M.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 680
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-680

Alignment Scores:

Pred. No.: 6 03e-152
Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%
Length: 665
Matches: 300
Conservative: 1
Mismatches: 1

Query Match:

DB: 95.39% Indels: 0
Gaps: 0

US-10-029-345a-108_COPY_538_1443 (1-906) x US-10-072-012-680 (1-665)

QY 1 ATGGCCCATGAGATGATTTGAACTCAAAATTGTTACTGAGAGATTGTGGCTCTGGAA 60
Db 1 MetAlaHisGluMetCileGlyThrGlnTleValThrGlnArgLeuValAlaLeuLeuGlu 20
QY 61 AGTGAACCGGAAAAAGTGTGCTATTGATGATGCCGCGCATTTGGGAAATTCATATCATCC 120
Db 21 SerGlyThrGlnValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
QY 121 CACATTTTGAAGCATTAAATATCAATCTCCCAAGCTTATGAAGCAAGTTCGCAACG 180
Db 41 HisLeuGlnValAlaLeuHisLeuHisLeuHisLeuHisLeuHisLeuHisLeuHisLeu 60
QY 181 GACAAAGCTTAAATTAACAAGCTCATTCAGCTTCAGCAATTCAGCAATTCAGCTTCAG 240
Db 61 AspLysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
QY 241 TGCAGTCAGAAAGTTGATGATTTACGATCAAACTCCCAAGATTTGCTCTCTCTTCA 300
Db 81 CysSerGlnLysValValValValValValValValValValValValValValValValVal 100
QY 301 GACTGTTTCTCACTGATCTTCTGCGTAACTCGAGAAAGCTTCACCTGCTTCACTG 360
Db 101 AspCysPheLeuThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
QY 361 CTTCGAGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
Db 121 LeuAlaGlyGlyPheAlaGlyPheSerArgCysPheProLysLeuCysGlnGlyLysSer 140
QY 421 ACTGATGCTTCACTGATTTTCACTGCTTCACTGCTTCACTGCTTCACTGCTTCACTG 480
Db 141 ThrLeuValProThrCysLeuSerGlnProCysLeuProValAlaAlaSerLeuLeuPro 160
QY 481 CGAATTCCTCCCAATCTTAACTTGTGGCTGCAGCGAGATGCTCAACAAGAGCTGATA 540
Db 161 ArgLeuLeuProAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
QY 541 CAGCAGATGGGATGGTATGTTGTTAAATGCCAGCTTACCTGTCCAAAGCTGACTTT 600
Db 181 GlnGlnAsnGlyIleGlyTyrrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAAGCTCAATTTCTGCGGTGCTGCTGCAATGACAGCTTTTGGAGAAATTTTG 660
Db 201 IleProLysThrHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
QY 661 CCGTGTGGACAAATCAGTATGATTTGATGAAAGCAAAAGCTTCCAAATGATGTGTT 720
Db 221 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTAGTCACTGTTTATGCTGGGATCTCCCGCTCCGCAATCTGCTTCAATCATCATG 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrrIleMet 260
QY 781 AAGAGAGTGAACATGCTTTAGATGAAGCTTACAGTTTGTGAAGAAAAAGAACTTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaTyrrArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCCAAATTCATTTTCTGGGCAACTCTTGAATGATGAGAAAGATTAAGAAC 900
Db 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrrGlnLysLysIleLysAsn 300
QY 901 CAGACT 906 |||||
Db 301 GlnThr 302

RESULT 5
US-10-168-506-14
; Sequence 14, Application US/10168506
; Publication No. US20040053229A1

```
/ GENERAL INFORMATION:
/ APPLICANT: PLOMMAN, GREGORY D.
/ APPLICANT: MARTINEZ, RICARDO
/ APPLICANT: WHITE, DAVID
/ APPLICANT: MANNING, GERARD
/ APPLICANT: SUDARSANAM, SUCHA
/ APPLICANT: HILL, RON
/ APPLICANT: PLANAGAN, PETER
/ TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
/ FILE REFERENCE: 038602/1351
/ CURRENT FILING DATE: 2002-06-21
/ PRIOR APPLICATION NUMBER: PCT/US00/34736
/ PRIOR FILING DATE: 2000-12-21
/ NUMBER OF SEQ ID NOS: 76
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 14
/ LENGTH: 665
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-168-506-14

Alignment Scores:
Pred. No.: 6, 03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 0
Query Match: 95.39% Indels: 0
Gaps: 0
DB: 12

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-168-506-14 (1-665)

QY 1 ATGGCCCATGAGATGATGGAATCAATGTTACTGAGAGGTTGGTCTGCTGGA 60
DB 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGGGAAGGAAAGAGTCTGCTTAATGATGAGCGGCACTTGTGGAATACATACATCC 120
DB 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
QY 121 CACATTTGGAGACCATTAATATCACTGCTCCAGCTTAATGAGGAAAGTTGCAACG 180
DB 41 HisIleLeuGluIleAlaIleLeuIleLeuIleLeuIleLeuIleLeuIleLeu 60
QY 181 GACAAAGTGTAAATTCACAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 240
DB 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAsp 80
QY 241 TGCAGTCAGAGGTTGATGATGATGATCAAGCTCCAGATGTTGCTCTCTCTCA 300
DB 81 CysSerGlnLysValValValValValValValValValValValValValValVal 100
QY 301 GACTGTTTCTCACTGCTACTCTGGGTAAACTGGAGAGAGCTTCACTCTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
QY 361 CTTCGAGGTGGGTTTCTGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlySerGluGlyLysSer 140
QY 421 ACTGTAGTCCCTACCTGAGCTTCTGAGCTGCTGTTACTGTTCCAAACATGGGCAACC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTCCTCCCAATCTTATCTTGGCTGCGACGAGATGTCTCTCAACAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuThrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuMet 180
QY 541 CAGCAGAAATGGGATTTGGTTATGTTTAATGCCAGATTAATCTGTTCAAGCTGACTTT 600
DB 181 GlnGlnAsnGlyIleGlyLysValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
```

```
DB 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
QY 661 CCGTGGTTGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCATGATGTGTT 720
DB 221 ProThrLeuAspLysSerValAspPheIleGlyLysAlaLysLysSerAsnGlyCysVal 240
QY 721 CTAGTGCACGTGTTTGGCTGGAGATCTCCGCTCCGACCATGCTGCTTGCCTACATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaThrIleMet 260
QY 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGTGAGAAAGCAAGCTTACT 840
DB 261 LysArgMetAspMetSerLeuAsnArgGluAlaThrArgPheValLysGluLysArgProThr 280
QY 841 ATATCTCAAATCTTCAATTTTCTGGGCCCACTCTTGATTAAGAGAAAGATTAAAGAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
QY 901 CAGACT 906
DB 301 GlnThr 302

RESULT 6
US-10-343-357-7
/ Sequence 7, Application US/10343357
/ Publication No. US20040058341A1
/ GENERAL INFORMATION:
/ APPLICANT: INCYTE GENOMICS, INC.; TANG, Y.Tom
/ APPLICANT: ELIOTT, Vlcki S.; RAMKUMAR, Jayalaxmi
/ APPLICANT: YAO, Monique G.; BURFORD, Neil
/ APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
/ APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
/ APPLICANT: LEE, Ernestine A.; HAPALIA, April J.A.
/ APPLICANT: LU, Dying Alina M.; TRIBOLEY, Catherine M.
/ APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
/ APPLICANT: YUE, Henry; WARREN, Bridget A.
/ APPLICANT: NGUYEN, Daniel B.; CHAMLA, Nandinder K.
/ APPLICANT: KEARNEY, Liam
/ TITLE OF INVENTION: PROTEIN PHOSPHATASES
/ FILE REFERENCE: PI-0173 PCT
/ CURRENT APPLICATION NUMBER: US/10/343,357
/ PRIOR FILING DATE: 2003-01-28
/ PRIOR APPLICATION NUMBER: PCT/US01/23716
/ PRIOR FILING DATE: 2001-07-26
/ PRIOR APPLICATION NUMBER: US 60/221,679
/ PRIOR FILING DATE: 2000-07-28
/ PRIOR APPLICATION NUMBER: US 60/223,272
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: US 60/224,309
/ PRIOR FILING DATE: 2000-08-10
/ PRIOR APPLICATION NUMBER: US 60/226,728
/ PRIOR FILING DATE: 2000-08-18
/ PRIOR APPLICATION NUMBER: US 60/229,254
/ PRIOR FILING DATE: 2000-08-30
/ PRIOR APPLICATION NUMBER: US 60/231,366
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PERL Program
/ SEQ ID NO 7
/ LENGTH: 665
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CD1
US-10-343-357-7

Alignment Scores:
Pred. No.: 6, 03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
```

Query Match:

DB: 95.39% Indels: 0
12 Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-377-072-26 (1-665)

```
QY 1 ATGGCCCATGAGATGATGGAATCTAAATGTTACTGAGAGTTGGTGGCTCTGCTGGA 60
DB 1 MetalaHisgluMetilegIythrGlnileValThrGluArgLeuValAlaLeuLeuGlu 20
QY 61 AGTGAACGGAAAAAGTGGCTGCTAATGATGACCGGCGCATTTGTGAAATCAATACATCC 120
DB 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
QY 121 CACATTTGGAGCCATTAATATCACTGCTCCAGCTTGAAGGAGGAGTTGCAACAG 180
DB 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIleuMetIleuValArgLeuGlnGln 60
QY 181 GACAAAGTGTATTAATACAGAGCTCATCAGCATTCAGCGAAACATAGAGTTGACATGAT 240
DB 61 AspLysValIleuIleThrGluLeuIleGlnHisSerIleValHisIleValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATGATTCAGTCAAAAGCTCCAGAGATGTTGCTCTCTCTTCA 300
DB 81 CysSerGlnLysValValValValValValValValValValValValValValValVal 300
QY 301 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGAGCTTCACTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTGTTCCCTGCGCTCTGGAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlySerGlyLeu 140
QY 421 ACTGATGTCCTCACTGATTTCTCAGCTTGTCTTACCTGTCGCAACATTTGGGCCAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 GCAATTTCTCCCAATCTTATCTTGGCTGCGCAGGAGATGCTCTCAACAGAGCTGATA 540
DB 161 ArgIleLeuProAsnLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 180
QY 541 CAGAGAAATGGAGATGTTGTTATGTTTAAATGACGCTAATACCTGTCGCAACCTGACTT 600
DB 181 GlnGlnAsnGlyIleGlyIleValIleuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTGCGTGGCTGCGTGAATGACAGCTTTGTGAGAAATTTTG 660
DB 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
QY 661 CCGTGGTTGGCAAAATCAGTATGTTCAATTGAGAAAGCAAAAGCTCCCATGATGTTGTT 720
DB 221 ProThrPheAspLysSerValAspPheIleGlyLysAlaLysAlaSerHisGlyCysVal 240
QY 721 CTATGTCAGCTGTTAGCTGGGATCTCCGCTCCGCAACATCCGCTATGCTATCATCATG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleuMet 260
QY 781 AAGAGATGACATGTTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTCT 840
DB 261 LysArgMetAspMetSerLeuAspGluAlaIleArgPheValIleuLysLysArgProThr 280
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGTCGATGAGAAAGAAAGATTAAGAAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIleGluLysIleLysAsn 300
QY 901 CAGACT 906
DB 301 GlnThr 302
```

RESULT 7

US-10-377-072-26

Sequence 26 Application US/10377072

Publication No. US20040009501A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Millennium Pharmaceuticals Inc.
/ APPLICANT: Curtis, Rory A.J.
/ APPLICANT: Logan, Thomas Joseph
/ APPLICANT: Glucksmann, Maria A.
/ APPLICANT: Meyers, Rachel E.
/ APPLICANT: Williamson, Mark J.
/ APPLICANT: Rudolph-Owen, Laura A.
/ APPLICANT: Tsai, Ping-Ying
/ APPLICANT: Tsai, Ping-Ying
/ TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117
/ TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
/ FILE REFERENCE: MP03-0180NM
/ CURRENT APPLICATION NUMBER: US/10/377,072
/ PRIOR FILING DATE: 2003-02-27
/ PRIOR APPLICATION NUMBER: US 09/895,860
/ PRIOR FILING DATE: 2001-06-29
/ PRIOR APPLICATION NUMBER: US 60/215,370
/ PRIOR FILING DATE: 2000-06-29
/ PRIOR APPLICATION NUMBER: US 60/187,455
/ PRIOR FILING DATE: 2000-03-07
/ PRIOR APPLICATION NUMBER: US 60/199,801
/ PRIOR FILING DATE: 2000-04-26
/ PRIOR APPLICATION NUMBER: US 60/199,801
/ PRIOR FILING DATE: 2001-05-21
/ PRIOR APPLICATION NUMBER: US 60/205,508
/ PRIOR FILING DATE: 2000-05-19
/ PRIOR APPLICATION NUMBER: US 09/816,494
/ PRIOR FILING DATE: 2001-03-23
/ PRIOR APPLICATION NUMBER: US 09/815,419
/ REMAINING PRIOR APPLICATION DATA REMOVED - See File Wrapper or PAM.
/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO: 26
/ LENGTH: 665
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-10-377-072-26
Alignment Scores:
Pred. No.: 5,03e-152
Score: 1552.00 Length: 665
Percent Similarity: 99.67% Matches: 300
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 1
DB: 15 Gaps: 0
US-10-029-345a-108_copy_538_1443 (1-906) x US-10-377-072-26 (1-665)
QY 1 ATGGCCCATGAGATGATGGAATCTAAATGTTACTGAGAGTTGGTGGCTCTGCTGGA 60
DB 1 MetalaHisgluMetilegIythrGlnileValThrGluArgLeuValAlaLeuLeuGlu 20
QY 61 AGTGAACGGAAAAAGTGGCTGCTAATGATGACCGGCGCATTTGTGAAATCAATACATCC 120
DB 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
QY 121 CACATTTGGAGCCATTAATATCACTGCTCCAGCTTATGAGAGGAGGAGTTGCAACAG 180
DB 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIleuMetIleuValArgLeuGlnGln 60
QY 181 GACAAAGTGTATTAATACAGAGCTCATCAGCATTCAGCGAAACATAGAGTTGACATGAT 240
DB 61 AspLysValIleuIleThrGluLeuIleGlnHisSerIleValHisIleValAspIleAsp 80
QY 241 TGCAGTCAGAGGTTGATGATTCAGTCAAAAGCTCCAGAGATGTTGCTCTCTCTTCA 300
```

Db 81 CySerGlnIuValValIyAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTTCTCACTGACTTCTGCGGTAACCTGGAGAAGCTTCAACTCTGTTCACTG 360
Db 101 AspCyPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAnsSerValHisLeu 120
QY 361 CTTCGAGGTGGGTTTGCTGAGTTCTCGTTGTTTCCCTGCGCTCTGAGGAAAATCC 420
Db 121 LeuValGlyValPheAlaGlnPheSerArgCyPheProGlyLeuGlyGlnLysSer 140
QY 421 ACTTAGTCCCTACCTGCACTTCTGAGCTTCTGCTTACTGCTGCAACATGGAGCCCAACC 480
Db 141 ThrLeuValProThrCysHisSerGlnProCysLeuProValAlaAsnHisLeuProThr 160
QY 481 CGAATTCCTCCCAATCTTTATCTTGCTGCTGCGCAGCAGAGATGCTCTCAACAAGAGCTGATA 540
Db 161 ArgHisLeuProAsnLeuIyLeuGlyCyGlnArgAspValLeuAsnLysGlnLeuMet 180
QY 541 CACGAGATGGAGTTGGTTATGTTTAAATGCCAGCTATACCTGTCGAAGCTGACTTT 600
Db 181 GlnGlnAsnGlyHisLeuGlyValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGCTGCTGCTGTAATGACAGCTTTGTGAGAAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysHisLeu 220
QY 661 CCGTGGTTGCAAAATCAGTAGATTTTATTGAGAAAAGCAAAAGCTTCAATGATGTTT 720
Db 221 ProTrpLeuAspLysSerValAspPheHisLeuLysAlaLysAsnHisLeuGlyCysVal 240
QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCCACATGCTATGCTCATCATCANG 780
Db 241 LeuValHisCysLeuAlaGlyHisLeuArgSerHisLeuAlaHisLeuAlaHisLeuMet 260
QY 781 AAGAGATGAGACATGCTTAGATGAAGCTTACAGATTTGTGAAGAAAAGACCTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaIyArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCCAAACTTCAATTTTCTGCGGCAACTCTGAGCTATGAGAAAAGATTAAAGAAC 900
Db 281 HisSerProAsnPheAsnPheLeuGlnLeuLeuAspTyGlnLysLysHisLeuAsn 300
QY 901 CAGACT 906
Db 301 GlnThr 302

RESULT 8

US-10-257-026-2
; Sequence 2, Application US/10257026
; Publication No. US20040086859A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10XDM5
; CURRENT APPLICATION NUMBER: US/10/257,026
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-257-026-2

Alignment Scores:

Pred. No.: 6,03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 0
DB: 16 Gaps: 0

US-10-029-345a-108_COPY_538_1443 (1-906) x US-10-257-026-2 (1-665)

QY 1 ATGCCCATGAGATGATGGAACTCAATTTGTTACTGAGAGGTTGGGCTGCTGAGAA 60
Db 1 MetAlaHisGlnMetHisLeuGlyHisGlnHisLeuAlaThrGlnArgLeuValAlaLeuLeuGln 20
QY 61 AGTGAACGGAAAATGCTGCTTAATGATAGCCGGCAATTTGTGAAATACAAATACATCC 120
Db 21 SerGlyThrGlnLysValLeuLeuHisLeuAspSerArgProPheValGlnIyArgHisSer 40
QY 121 CACATTTGGAAACCATTAATATCAATGCTCCAGCTTATGAAAGGAAAGTTGCAACAG 180
Db 41 HisHisLeuGlnAlaHisLeuHisLeuHisLeuHisLeuHisLeuHisLeuHisLeuHisLeu 60
QY 181 GACAAGGTGTAATTCACAGCTCAATCCAGCATTCACCGAAATGAATGAAGTTGACATTTAT 240
Db 61 AspLysValLeuHisLeuThrGlnLeuHisLeuHisLeuHisLeuHisLeuHisLeuHisLeu 80
QY 241 TGCAGTCAGAAGGTTTATGATTACGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
Db 81 CySerGlnIuValValIyAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACTGTTTTCTCACTGACTTCTGCGGTAACCTGGAGAAGCTTCAACTCTGTTCACTG 360
Db 101 AspCyPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAnsSerValHisLeu 120
QY 361 CTTCGAGGTGGGTTTGCTGAGTTCTCGTTGTTTCCCTGCGCTCTGAGGAAAATCC 420
Db 121 LeuValGlyValPheAlaGlnPheSerArgCyPheProGlyLeuGlyGlnLysSer 140
QY 421 ACTTAGTCCCTACCTGCACTTCTGAGCTTCTGCTTACTGCTGCAACATGGAGCCCAACC 480
Db 141 ThrLeuValProThrCysHisSerGlnProCysLeuProValAlaAsnHisLeuProThr 160
QY 481 CGAATTCCTCCCAATCTTTATCTTGCTGCTGCGCAGCAGAGATGCTCTCAACAAGAGCTGATA 540
Db 161 ArgHisLeuProAsnLeuIyLeuGlyCyGlnArgAspValLeuAsnLysGlnLeuMet 180
QY 541 CACGAGATGGAGTTGGTTATGTTTAAATGCCAGCTATACCTGTCGAAGCTGACTTT 600
Db 181 GlnGlnAsnGlyHisLeuGlyValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGCTGCTGCTGTAATGACAGCTTTGTGAGAAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysHisLeu 220
QY 661 CCGTGGTTGCAAAATCAGTAGATTTTATTGAGAAAAGCAAAAGCTTCAATGATGTTT 720
Db 221 ProTrpLeuAspLysSerValAspPheHisLeuLysAlaLysAsnHisLeuGlyCysVal 240
QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCCACATGCTATGCTCATCATCANG 780
Db 241 LeuValHisCysLeuAlaGlyHisLeuArgSerHisLeuAlaHisLeuAlaHisLeuMet 260
QY 781 AAGAGATGAGACATGCTTAGATGAAGCTTACAGATTTGTGAAGAAAAGACCTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaIyArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCCAAACTTCAATTTTCTGCGGCAACTCTGAGCTATGAGAAAAGATTAAAGAAC 900
Db 281 HisSerProAsnPheAsnPheLeuGlnLeuLeuAspTyGlnLysLysHisLeuAsn 300
QY 901 CAGACT 906
Db 301 GlnThr 302

RESULT 9

US-10-648-593-240
; Sequence 240, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR

```

; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 240
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-240

Alignment Scores:
Pred. No.: 6 03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 0
Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-648-593-240 (1-665)

QY 1 ATGGCCCATGAGTGAATTGAACTCAAAATTGTTACTGAGAGGTGGTGGCTGCTGGAA 60
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAAAGTGCTGCTATGATGACCGGCGCATTTGTGAATACATACATCC 120
Db 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluYrAnThrSer 40
QY 121 CACATTTGGAGCCATTAATATCACTGCTCAAGCTTATGAAGGAAGTTGACACG 180
Db 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleuMetIysArgIleuGln 60
QY 181 GACAAAGTGTAAATTAACAGCTCATCCAGCATTCAGCCAAACATAAAGTTGACATTGAT 240
Db 61 AspIysValIleuIleThrGlnIleuIleGlnHisSerAlaIysIleValAspIleAsp 80
QY 241 TGCAGTCGAAGGTGTGATAGTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 300
Db 81 CysSerGlnIysValValValIyrAspGlnSerSerGlnAspValAlaSerIleuSer 100
QY 301 GACTGTTTTCACCTGACTCTGGGTAACATGGAAGAGACTTCACTCTGTTCACTG 360
Db 101 AspCysPheLeuThrValIleuLeuGlyIysLeuGluIysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGGTGGTTGCTGATCTCTGTTGTTCCCTGGCTCTGTGAAGAAATCC 420
Db 121 LeuAlaGlyValPheAlaGlnPheSerArgCysPheProGlyLeuGlyGluGlyIysSer 140
QY 421 ACTCTGCTCTACCTGCACTTCTAGGCTTGCTTACCTGTTGGCAACATTTGGGCCAAC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATCTCTCCCAACTTTATCTTGGTGGCCAGGAGATGCTTCAACAAGAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuIyrLeuGlyCysGlnArgAspValIleuAsnIysGluLeuMet 180
QY 541 CAGCAGATGGATGGTATGTTATGTTAAAGCCAGCATATCCCTGTCAAAGCTTACTTT 600
Db 181 GlnGlnIleuGlnIleGlyIyrValIleuAsnAlaSerAsnThrCysProIysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCTCGCGTGGCTGTGATGACAGCTTTTGTGAAGAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIysIleLeu 220
QY 661 CCGTGTGTGAACAATCATAGATTTTCATTGGAAGAAAGAGCTCCAAATGAGATGGTT 720
Db 221 ProIyrPheuAspIysSerValAspPheIleGluIysAlaIysAlaSerAsnGlyCysVal 240
QY 721 CTAGTGACATGTTTACTGTGGATCTCCGCTCCGACCAATCGCTATGCTCATCATCATG 780

; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 247
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-247

Alignment Scores:
Pred. No.: 6 03e-152 Length: 665
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 0
Gaps: 0

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-648-593-247 (1-665)

QY 1 ATGGCCCATGAGTGAATTGAACTCAAAATTGTTACTGAGAGGTGGTGGCTGCTGGAA 60
Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
QY 61 AGTGAACGGAAGAAAGTGCTGCTATGATGACCGGCGCATTTGTGAATACATACATCC 120
Db 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluYrAnThrSer 40
QY 121 CACATTTGGAGCCATTAATATCACTGCTCAAGCTTATGAAGGAAGTTGACATTGAT 240
Db 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleuMetIysArgIleuGln 60
QY 181 GACAAAGTGTAAATTAACAGCTCATCCAGCATTCAGCCAAACATAAAGTTGACATTGAT 240
Db 61 AspIysValIleuIleThrGlnIleuIleGlnHisSerAlaIysIleValAspIleAsp 80
QY 241 TGCAGTCGAAGGTGTGATAGTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCTCA 300
Db 81 CysSerGlnIysValValValIyrAspGlnSerSerGlnAspValAlaSerIleuSer 100
QY 301 GACTGTTTTCACCTGACTCTGGGTAACATGGAAGAGACTTCACTCTGTTCACTG 360
Db 101 AspCysPheLeuThrValIleuLeuGlyIysLeuGluIysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGGTGGTTGCTGATCTCTGTTGTTCCCTGGCTCTGTGAAGAAATCC 420

```

Db	121	LeuAlaGlyValYpheaIaGluPheSerAxCysPheProGlyLeuCYaGluGlyLysSer	140
Qy	421	ACTCTAGTCCCTACCTGCATTTTCTACAGCTTGTCTTACTGTGGCCAAATTTGGCCAAAC	480
Db	141	ThIleValaProThrCysIleSerGlnProCysLeuProValAlaIleuIleGlyProThr	160
Qy	481	CGAATCTTCCCAATCTTTATCTGTGGCTGCGACGAGATGTCCTCAACAAAGACTGATA	540
Db	161	ArgIleuProAsnLeuTyrLeuGlyCysGlnArgPheValIleuAsnIleGluIleuMet	180
Qy	541	CACGAGATGGATGGTATTGTATTGTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT	600
Db	181	GlnGlnAsnGlyIleGlyTyrValIleuAsnAlaSerAsnThrCysProIysProAspPhe	200
Qy	601	ATCCCCGAGTCTTATTTCTCGCTGTGCTGTGATGACAGCTTTTGTGAGAAAATTTTG	660
Db	201	IleProGluSerHisPheLeuArgValProValIleuAspSerPheCysGluIysIleLeu	220
Qy	661	CGCTGGTGTGACAAATCAGTAGATTTCATTGAGAAAGAAAGCCTCCATAGATGTGTT	720
Db	221	ProTyrLeuAspLysSerValAspPheIleGluIysAlaIysPheAlaSerAsnGlyCysVal	240
Qy	721	CTAGTGCACACTTTTACTGCGGATCTCCGCTCCGACACATCGCTATCGCTATCATATG	780
Db	241	LeuValaHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleTyrIleMet	260
Qy	781	AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAGACTTACT	840
Db	261	LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValIysGluIysArgProThr	280
Qy	841	ATATCTCCAACTTCAATTTTCTGGGCGCAACTCTCTGACTATGAGAGAAATTAAGAAC	900
Db	281	IleSerProAsnPheAsnIleLeuGlyIleuLeuAspTyrGluIysIleIysAsn	300
Qy	901	CAGACT 906	
Db	301	GlnThr 302	
RESULT 11			
US-10-072-012-679			
: Sequence 679, Application US/10072012			
: Publication No. US20040033493A1			
: GENERAL INFORMATION:			
: APPLICANT: Tchernev, Velizar			
: APPLICANT: Spytek, Kimberly			
: APPLICANT: Zerhusen, Bryan			
: APPLICANT: Patturajan, Meera			
: APPLICANT: Shimkete, Richard			
: APPLICANT: Li, Li			
: APPLICANT: Gangolli, Esha			
: APPLICANT: Padigaru, Muralidhara			
: APPLICANT: Anderson, David W.			
: APPLICANT: Rastelli, Luca			
: APPLICANT: Miller, Charles E.			
: APPLICANT: Gerlach, Valerie			
: APPLICANT: Taupier Jr, Raymond J.			
: APPLICANT: Gusev, Vladimir Y.			
: APPLICANT: Coleman, Steven D.			
: APPLICANT: Wolenc, Adam R.			
: APPLICANT: Pena, Carol E. A			
: APPLICANT: Furtak, Katarzyna			
: APPLICANT: Grose, William M.			
: APPLICANT: Alsobrook Jr, John P.			
: APPLICANT: Lepley, Denise M.			
: APPLICANT: Rieger, Daniel K.			
: APPLICANT: Burgess, Catherine E.			
: TITLE OR INVENTION: Proteins and Nucleic Acids Encoding Same			
: FILE REFERENCE: 21402-258			
: CURRENT APPLICATION NUMBER: US/10/072,012			
: PRIOR FILING DATE: 2002-01-31			
: PRIOR APPLICATION NUMBER: 60/265,102			
: PRIOR FILING DATE: 2001-01-30			
: PRIOR APPLICATION NUMBER: 60/265,514			

[illegible]

Db 206 GINGINaenGlyIleGlyTyValleuAnaAlaSerAenThnCysProIysProAspSph 225
Qy 601 ATCCCCGAGCTTCATTTTCTGCGCTGTGCTCCGTGAATGACAGCTTTTGTGAGAAAAATTTCG 660
Db 226 ILePProGluSerHisSphLeuArgValProValAsnaSerPheCysGluIleGlyLeu 245
Qy 661 CCGTGGTTGAGCAAAATCAGATGATTCATTGATGAGAAAGCAAAAGCCTCCAAATGAGATGTGTT 720
Db 246 ProIlePuenAplysSerValaIaSPheIleGluIySaIaIySaIaSerHisnGlyCysVal 265
Qy 721 CTAGTGCACTGTTTAGCTGGGATCTCCGCTCCGACACCATCGCTATCGCTACATCATG 780
Db 266 LeuValHisCysLeuAlaGlyIleSerHisSerAlaTrIleAlaIleAlaTyIleMet 285
Qy 781 AAGAGATGTGACATGTCTTTAGATGAAGCTTAACAGATTGTGAAAGAAAAAACACTACT 840
Db 286 LysArgMetCysPheCysSerIeuaSpGluAlaTyIrnArgPheValIySgluIySaIaGProThr 305
Qy 841 ATATCTCCAAACTTCATTTTCTGGGCCAACTCTCGCATATGATGAGAGAAGATTAAGAAC 900
Db 306 ILeSerProAsnPhaSnPheLeuGlyCylLeuLeuaSpTyGluIySylsIleIySaSn 325
Qy 901 CAGACT 906
Db 326 GlnThr 327

RESULT 12
 US-10-072-012-703 / Sequence 703, Application US/10072012
 Publication No. US20040033493A1
 GENERAL INFORMATION:
 APPLICANT: Tchernev, Velizar
 APPLICANT: Spytek, Kimberly
 APPLICANT: Zernhusen, Bryan
 APPLICANT: Patturajan, Meera
 APPLICANT: Shinkets, Richard
 APPLICANT: Li, Li
 APPLICANT: Gargolli, Esha
 APPLICANT: Padigaru, Muraidhara
 APPLICANT: Anderson, David W.
 APPLICANT: Rastelli, Luca
 APPLICANT: Miller, Charles E.
 APPLICANT: Gerlach, Valerie E.
 APPLICANT: Taupier, Jr, Raymond J.
 APPLICANT: Gusev, Vladimir Y.
 APPLICANT: Colman, Steven D.
 APPLICANT: Wolenc, Adam R.
 APPLICANT: Pena, Carol E. A
 APPLICANT: Furtak, Katarzyna
 APPLICANT: Grosse, William
 APPLICANT: Alsobrook II, John P.
 APPLICANT: Lepley, Denise M.
 APPLICANT: Rieger, Daniel K.
 APPLICANT: Burgess, Catherine E.
 TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 FILE REFERENCE: 21402-258
 CURRENT APPLICATION NUMBER: US/10/072,012
 CURRENT FILING DATE: 2002-01-31
 PRIOR APPLICATION NUMBER: 60/265,102
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: 60/265,514
 PRIOR FILING DATE: 2001-01-31
 PRIOR APPLICATION NUMBER: 60/265,517
 PRIOR FILING DATE: 2001-01-31
 PRIOR APPLICATION NUMBER: 60/265,412
 PRIOR FILING DATE: 2001-01-31
 PRIOR APPLICATION NUMBER: 60/265,395
 PRIOR FILING DATE: 2001-01-31
 PRIOR APPLICATION NUMBER: 60/266,406
 PRIOR FILING DATE: 2001-02-02
 PRIOR APPLICATION NUMBER: 60/266,767
 PRIOR FILING DATE: 2001-02-05

```

? PRIOR APPLICATION NUMBER: 60/267,057
? PRIOR FILING DATE: 2001-02-07
? PRIOR APPLICATION NUMBER: 60/266,975
? PRIOR FILING DATE: 2001-02-07
? PRIOR APPLICATION NUMBER: 60/267,459
? PRIOR FILING DATE: 2001-02-08
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 1391
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 703
? LENGTH: 690
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-072-012-703

```

Alignment Scores:	
Pred. No.:	
Score:	6.09e-152
Percent Similarity:	1552.00
Best Local Similarity:	99.67%
Query Match:	99.34%
DB:	95.39%
	12
	Gaps:
	0
	Length:
	690
	Matches:
	300
	Conservative:
	1
	Mismatches:
	1
	Indels:
	0
	Gaps:
	0

US-10-029-345A-108_COPY_538_1443 (1-906) X US-10-072-012-703 (1-690)

QY	1	ATGGCCCATGAGATGATGGAACTCAATTTGTTACTGAGAGTTGGCTGTCTGTGGAA	60
Db	26	MetAahsGumectIiegIythrInIeValThrtGutrgIeuValAlaIeuIeuGlu	45
QY	61	AGTGAACGGAAGAAAGCTGCTAAATGATGCGGCATTGTGGAAATACATCATCC	120
Db	46	SerGlythrGluIuysValIeuIeuIeAspSerXrProPheValIgluYrAsnThrSer	65
QY	121	CACATTTTGGAAACCATTAATATCACTGCTCCAGGCTTATGAAAGGAAAGTTGACAG	180
Db	66	HisIleuendIuAlaIeAsnIleAsnCySerIyIeuMelIyAhrGluengIIn	85
QY	181	GACAAAGTTTAATTAACAGACTCATCATTCACATTCGCAACATTAAGTTTGACATTGAT	240
Db	86	AspIyValIeuIleThrtGluIeuIeGlnHisSerIalYshIsIyValAspIleAsp	105
QY	241	TGAGTCACAAGCTTGATGTTTACGATCAAAAGCTCCAGAGTGTGCTCTCTCTCA	300
Db	106	CysSerGluYsValIeValYIyrAspGInSerSerGlnAspValAlaSerIeAspSer	125
QY	301	GACTGTTTCTCACTGACTTCTGGGTAAACTGAGAAAGACTTCAACTCTGTCACTCG	360
Db	126	AspCySphIeuThrValIeuIeuGluYIyIeuGluYIySerPheAsnSerValIhIsIeu	145
QY	361	CTTGAGGTGGGTTTGGTGAGTCTCTGTTTCTCCGTGCTGTGAGAGAAATCC	420
Db	146	IeuAlaGluYglYpheaIeGluPheSerXrYcysPheProGluYeuGluYIySer	165
QY	421	ACTGACTGCCCTACCTGCAATTTCTAGAGCTTGCTTAACTGTGCGCAACTGGGCCAAC	480
Db	166	ThrIeuValProThrCysIleSerGlnProCysIeuProValAlaAsnIleIyProThr	185
QY	481	GGAATTCCTCCAAATCTTATCTTGGCGTCAGCGAGATGTCTCAACAAGAGCTGATA	540
Db	186	ArgIleIeuProAsnIeuYrIeuGluYcysGlnAspValIeuAsnIySGIuIeuMet	205
QY	541	CAGAGAAATGGGATTTGGTTATGTGTTAAATCCAGATATACTGTCGCAAACTGCACTTT	600
Db	206	GlnGlnAsnGluYIeGluYrValIeuAsnAlaSerAsnThrCysProIySproAspPhe	225
QY	601	ATCCCGAAGTCAATTCCTGCGTGGCCGTGGAAATGACGTTTGTGAGAAATTTTG	660
Db	226	IleProGluIeuSerHisPheIeuAsrValProValAsnAspSerPheCysGluYsIleIeu	245
QY	661	CCGTGGTTTGAACAATAGTAAATTCATTATGAGAAAGCAAAAGCTCCATATGATGTGT	720
Db	246	ProIleuAspIySerValAspPheIleGluYsAlaIyAlaSerAsnGluYcysVal	265

```

OY 721 TAGGCACTGTTTAGCTGGAGATCTCCCGCTCCGACATGCGCTTGCCTACATCAG 780
Db 266 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleLeu 285
OY 781 AAGAGATGATGACATGCTCTTTAGATGAAAGCTTACAGATTGTGAAAGAAAAAGACTTACT 840
Db 286 LysAlaGlyMetAspMetSerLeuAspGlnAlaTyrArgPheValIleSerIuysArgP 305
OY 841 ATATCTCCAAACTTCAATTTTCTGGCCCAACTCTCGGACTATGAGAAAGAAATTAAAGAAC 900
Db 306 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluIuysIleIysAsn 325
OY 901 CAGACT 906
Db 326 GlnThr 327

RESULT 13
US-10-425-114-54204
: Sequence 54204, Application US/10425114
: Publication No. US20040034888A1
: GENERAL INFORMATION:
: APPLICANT: Liu, Jindong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovalic, David K.
: APPLICANT: Screen, Steven E
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Use Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53)13B
: CURRENT APPLICATION NUMBER: US/10/425,114
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 54204
: LENGTH: 690
: TYPE: PRT
: ORGANISM: Homo sapiens
: OTHER INFORMATION: Clone ID: LIB4119-028-H6_PL1.pep
US-10-425-114-54204

Alignment Scores:
Pred. No.: 6,09e-152 Length: 690
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 95.39% Indels: 0
DB: 12 Gaps: 0

US-10-029-345A-108_COBY_538_1443 (1-906) x US-10-425-114-54204 (1-690)
OY 1 ATGGCCCCATGAGATGATTGGAACCTCAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGAA 60
Db 26 MetAlaHisGluMetIleGlyThrGlnIleValIleThrGluArgLeuValAlaLeuLeuGln 45
OY 61 AGTGAACGGAAGAAAGTCTGCTGAATTGATACCGCGCCATTTGTGGAATACATACATCC 120
Db 46 SerGlyThrGluIuysValLeuLeuIleAspSerArgProPheValGluThrAsnThrSer 65
OY 121 CACATTTTGGAAAGCATTAATATATCAACTGCTTCAAGCTTATGGAAGCGAAGTTTGCAACG 180
Db 66 HisIleLeuGlnAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 85
OY 181 GACAAAGGTGTAATATACAGAGCTCATCCAGCACTTCAGCAACATTAAGGTTGACATTGAT 240
Db 86 AspIysValIleLeuIleThrGlnLeuIleGlnHisSerAlaIuysHisIleValAspIleAsp 105
OY 241 TGCAGTCGAAGAGTTGTAGATTACGATCAAAAGCTCCCAAGATGTTGCCCTCTCTCTTCA 300
Db 106 CysSerGlnIuysValIleValIuysArgPheGlnSerGlnAspValAlaSerLeuSerSer 125
OY 301 GACGTTTTCACCTGACTTCTGGGTAACTGAGAGAAGACTTCAACTCTGTTCACCTG 360

```

Db	126	AspCysPheIeuThrValIleuLeuGlyLysIeuGluLysSerPheAsnSerValHisIleu	145
QY	361	CTTGCAAGTGGGTTTGGTCTGAGCTTCTCTGCTGTTTCCCTGGCCCTCTGTGAAGAAATCC	420
Db	146	LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlySerGlyLysSer	165
QY	421	ACTGAGTCCCACTGCACTTCTGAGCTTGCCTTAACCTGTTGGCCAACTGGGCGCAACC	480
Db	166	ThrIleuValProThrCysHisIleSerGlnProCysIleuProValAlaIleuHisIleGlyProThr	185
QY	481	CGAATTTCTCCCAATCTTATCTTGGCTGCGCCGCGAGATGTCCTCAACAGAGAGCTGATA	540
Db	186	ArgIleIleuProIleuPheIleuThrIleuGlyCysGlnArgAspValIleuAsnIleGlyIleuMet	205
QY	541	CAGCAGAAATGGAGTGGTTATGTATGTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT	600
Db	206	GlnGlnAlaGlnGlyIleGlyIleValIleuAlaIleSerAsnThrCysProIleProIlePhe	225
QY	601	ATCCCCGAGTCTCACTTCCCTGCGTGCCTGGAATGACAGCTTTTGGAGAAATTTTG	660
Db	226	IleProGlnSerHisPheIleuArgValProValAsnAspSerPheCysGluLysIleu	245
QY	661	CCGTGGTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCCTCAATGGATGTGT	720
Db	246	ProThrIleuAspLysSerValAspPheIleGlyLysAlaIleAsnIleGlySerVal	265
QY	721	CTAGTGACTGTTAGTGGGATCTCCGCTCCGCGCACCATGCTATGCTTACATCATG	780
Db	266	LeuValHisCysIleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleThrIleMet	285
QY	781	AAGAGGATGAGACATCTCTTATAGATGAAGCTTACAAATTTGTGAAAGAAAGAAAGCTACT	840
Db	286	LysArgMetAspMetSerIleuAspGlnAlaIleArgPheValIleGlyLysArgProThr	305
QY	841	ATATCTCCAAACTCAATTTTCTGAGCCCACTCTGACATATGAGAAAGAAAGATTAGAAG	900
Db	306	IleSerProIleuPheAsnIleuGlyIleuIleuAspTyrGluLysIleuLysAsn	325
QY	901	CAGACT 906	
Db	326	GlnThr 327	

RESULT 14
 US-10-072-012-681
 ; Sequence 681, Application US/10072012
 ; Publication No. US20040033493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Therney, Velizar
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Zernusen, Bryan
 ; APPLICANT: Patuzaian, Meera
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Li, Li
 ; APPLICANT: Gangoli, Esha
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Anderson, David W.
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Taubier Jr, Raymond J.
 ; APPLICANT: Gusev, Vladimir Y.
 ; APPLICANT: Colman, Steven D.
 ; APPLICANT: Wolenc, Adam R.
 ; APPLICANT: Pena, Carol E. A
 ; APPLICANT: Furtak, Katarzyna
 ; APPLICANT: Grose, William M.
 ; APPLICANT: Alsobrook II, John P.
 ; APPLICANT: Lepley, Denise M.
 ; APPLICANT: Rieger, Daniel K.
 ; APPLICANT: Burgess, Catherine E.
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-258
 ; CURRENT APPLICATION NUMBER: US/10/072.012

Percent Similarity: 99.34%
Best Local Similarity: 99.01%
Query Match: 94.96%
DB: 15

Conservative: 1
Matches: 2
Indels: 0
Gaps: 0

Job time : 51.0982 secs

US-10-029-345a-108_copy_538_1443 (1-906) x US-10-094-749-2312 (1-665)

```
OY 1 ATGGCCATGATGATTGGAACTCAAAATGTTACTGAGAGATTGGTGGCTGCTGGAA 60
DB 1 MetAlaHisGlnMetCileGlyThrGlnIleValThrGlnArgLeuValAlaLeuGln 20
OY 61 AGTGAACGGAAAAAGTGTCTAATTGATAGCCGGCCATTGTGGAAATACAAATACAC 120
DB 21 SerGlyThrGlnValLeuIleAspSerArgProPheValGluTyraAsnThrSer 40
OY 121 CACATTTGGAAAGCCATTATATCACTGCTCCAAAGCTTATGAAGCAAGTTGCAACAG 180
DB 41 HisIleuGlnIuAlaIleAsnIleAsnCySerIlySerIlySerArgArgLeuGln 60
OY 181 GACAAAGTGTATTACAGAGCTCATCCAGCATTCAGCGAAACATAAGATTGAT 240
DB 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
OY 241 TGCAGTCAGAAAGTTGATTACGATCAAAAGCTCCCAAGATGTGCTCTCTTCA 300
DB 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerIleuSer 100
OY 301 GACTGTTTTCACCTGATCTTGGGGTAACTGAGAAAGCTTCAACTGTTCACTG 360
DB 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
OY 361 CTTCGAGGTGGTGGTGGTGGTCTCTCGTGTGTTCCCTGGCCCTGTGAAAGAAATCC 420
DB 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyLysSer 140
OY 421 ACTCTAGTCCCTACCTGCACTTCTCAGCCTTGCTTACCTGTTCCAAACTTGGGCCAAC 480
DB 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
OY 481 CGAATCTTCCCAATCTTTATCTTGGCTGCCAGCAGATGCTCTCAACAAAGAGCTGATA 540
DB 161 LeuIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGlnLeuMet 180
OY 541 CAGCAGAAATGGATGGTATGTTAAATGCCAGCATATACCTGTCAAAGCCTGACTT 600
DB 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLysProAspPhe 200
OY 601 ATCCCGAGTCTCATTTCTCGCGTGCCTGTGAATGACAGCTTTGTGAGAAATTTTG 660
DB 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
OY 661 CCGTGTGGACAAATGATGATTTTCATTGAGAAAGCAAAAGCCTCCAAATGGATGTT 720
DB 221 ProThrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
OY 721 CTAGTGCACGTGTTAGTGGGATCTCCCGCTCCGCCACATCGCTATCGCTTACATCAG 780
DB 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
OY 781 AAGAGATGACATGCTCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGCCTACT 840
DB 261 LysArgMetCysPheSerLeuAspGlnAlaTyrArgPheValLysGlnLysArgProThr 280
OY 841 ATATCTCCAAACTTCAATTTTGTGGGCCAACTCCTGGACTATGAGAAAGATTAAAGAC 900
DB 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysLysIleLysAsn 300
OY 901 CAGACT 906
DB 301 GlnThr 302
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:28:26 ; Search time 11.5553 Seconds

(without alignments)
8095.510 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 1627
Sequence: 1 atggccacatagatgatgttg.....agaagataaaccagact 906

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delpext 7.0	

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+np.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10029345/runat_21062004_122816_4233/app_query.fasta_1.3278
-DB=Issued Patents AA -QFMT=fastaan -SUFPIX=rai -MINMATCH=0.1 -LOOPLC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345_QCGN_1_1_45 @runat_21062004_122816_4233 -NCPU=6
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA.*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/1aa/5A_COMB.pep.*
4: /cgn2_6/prodata/2/1aa/5B_COMB.pep.*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1552	95.4	665	4	US-09-816-494-2
2	654.5	40.2	170	4	US-09-544-716-14
3	654.5	40.2	170	4	US-09-557-921-15
4	654.5	40.2	170	4	US-09-564-357-17
5	654.5	40.2	170	4	US-09-619-380-16
6	469	28.8	482	4	US-09-557-921-2
7	427	26.2	394	4	US-08-530-290-23
8	427	26.2	394	4	US-09-702-705-805
9	427	26.2	394	4	US-09-702-705-827
10	427	26.2	394	4	US-09-736-457-805
11	427	26.2	394	4	US-09-736-457-827
12	427	26.2	394	4	US-09-614-124B-805

13	427	26.2	394	4	US-09-614-124B-827	Sequence 827, App
14	427	26.2	394	4	US-09-671-325-805	Sequence 805, App
15	427	26.2	394	4	US-09-671-325-827	Sequence 827, App
16	427	26.2	394	4	US-09-589-184-805	Sequence 805, App
17	427	26.2	394	4	US-09-589-184-827	Sequence 827, App
18	425	26.1	314	4	US-09-371-671B-11	Sequence 11, App
19	425	26.1	367	2	US-08-990-379-6	Sequence 6, App1
20	423.5	26.0	395	2	US-08-990-379-5	Sequence 5, App1
21	415	25.5	367	2	US-08-530-290-24	Sequence 24, App1
22	409.5	25.2	313	2	US-08-990-379-7	Sequence 7, App1
23	409	25.1	314	3	US-09-164-193-22	Sequence 22, App1
24	409	25.1	314	4	US-09-921-448A-22	Sequence 22, App1
25	377.5	23.2	393	2	US-08-990-379-4	Sequence 4, App1
26	357	21.9	302	4	US-09-702-705-806	Sequence 806, App
27	357	21.9	302	4	US-09-736-457-806	Sequence 806, App
28	357	21.9	302	4	US-09-614-124B-806	Sequence 806, App
29	357	21.9	302	4	US-09-671-325-806	Sequence 806, App
30	357	21.9	302	4	US-09-589-184-806	Sequence 806, App
31	349.5	21.5	397	2	US-08-990-379-8	Sequence 8, App1
32	346.5	21.3	168	4	US-09-544-716-13	Sequence 13, App1
33	346.5	21.3	168	4	US-09-557-921-13	Sequence 13, App1
34	346.5	21.3	168	4	US-09-564-357-16	Sequence 16, App1
35	346.5	21.3	168	4	US-09-619-380-15	Sequence 15, App1
36	338.5	20.8	170	4	US-09-544-716-12	Sequence 12, App1
37	338.5	20.8	170	4	US-09-557-921-12	Sequence 12, App1
38	338.5	20.8	170	4	US-09-564-357-15	Sequence 15, App1
39	338.5	20.8	170	4	US-09-619-380-14	Sequence 14, App1
40	319	19.6	169	4	US-09-544-716-16	Sequence 16, App1
41	319	19.6	169	4	US-09-557-921-17	Sequence 17, App1
42	319	19.6	169	4	US-09-564-357-19	Sequence 19, App1
43	319	19.6	169	4	US-09-619-380-18	Sequence 18, App1
44	312	19.2	168	4	US-09-544-716-15	Sequence 15, App1
45	312	19.2	168	4	US-09-557-921-16	Sequence 16, App1

ALIGNMENTS

RESULT 1
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816, 494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-494-2

Alignment Scores:
Pred. No.: 3.93e-181
Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%
Query Match: 95.39%
DB: 4
Gaps: 0

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-816-494-2 (1-665)

QY 1 ATGGCCCATGAGATGATTTGCACTCAATTTGTTACTGAGAGCTTGCTGCTGCAAA 60
DB 1 Mettllahhlgllwettllgllthrglnllvalthghlartgllvllalaleuugll 20
QY 61 AGTGAACGGAAGAAAGTGTGCTGCTAATTGATTGATAGCGGCATTTGTGGAATACATCATCC 120

No.:	2.22e-71
e:	654 E0
Length:	170

Percent Similarity: 88.62% Conservative: 27
Best Local Similarity: 72.46% Mismatches: 18
Query Match: 40.23% Indels: 1
Gaps: 1
US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-557-921-15 (1-170)

QY 400 GGGCTGTGTGAAGAAAA--TCACCTTGTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTA 456
Db 1 G1yleuCy6glucllyysProAlaAlaLeuLeuProMeSerLeuSerGlnProCy6leu 20
QY 457 CCTGTGGCCAAACATTGGGCCCAACCCGAATTTCTCCCAATTTTATTTGCTGGCTCCAGCGA 516
Db 21 ProValProSerValG1yleuThrArg1leuProH1leuProH1leuGlySerGlnlys 40
QY 517 GATGTCTTCAACAAGAGCTGATATACAGCAAGATGGATGGTATATGTATGTTAAATGCCAGC 576
Db 41 AspValLeuAsnLysAspLeuMetThrGlnAsnGly1leSerTyValLeuAsnAlaSer 60
QY 577 TATACCTGTCCAAAGCCTGACTTATCCCGAGTCTCATTTCCGTGCTGCTGGAAT 636
Db 61 AsnSerCy6ProLysProAspPhe1leCy6GluSerArgPheMetArgValPro1leAsn 80
QY 637 GACAGCTTTGTGAGAAAAATTTGGCGTGGTGGCAAAATCAGATGATTGATTGAGAA 696
Db 81 AspAsnTyCy6gluLysLeuLeuProTyrLeuAspLysSer1leGluPhe1leAspLys 100
QY 697 GCAAAAGCCTCCAAATGATGTGTCTAGTCACTGTTAGCTGGATCTCCGCTCGCC 756
Db 101 AlAlyLeuSerSerCy6gluVal1leValH1sCy6leuAlaGly1leSerArgSerAla 120
QY 757 ACCATGCTATGCGCTACATCATGAGAGATGACATGCTTTAATGATGAAGCTTACAGA 816
Db 121 Thr1leAla1leAlaTyrlleMetLysThrMetGlyMeSerSerAspAlaTyArg 140
QY 817 TTTGTGAAGAAAAAGACCTACTATATCTCAAACTTCAATTTTCTGGGCCAACTCTG 876
Db 141 PheValLysAspArgArgProSer1leSerProAsnPheAsn1leuGlyGlnLeu 160
QY 877 GACTATGAGAAGATTAG 897
Db 161 GluTyrgluArgThrLeuLys 167

RESULT 4
US-09-564-357-17
; Sequence 17, Application US/09564357
; Patent No. 6645753
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; TITLE OF INVENTION: DSP-5 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.413
; CURRENT APPLICATION NUMBER: US/09/564,357
; CURRENT FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-564-357-17

Alignment Scores:
Pred. No.: 2,22e-71 Length: 170
Score: 654.50 Matches: 121
Percent Similarity: 88.62% Conserves: 27
Best Local Similarity: 72.46% Mismatches: 18
Query Match: 40.23% Indels: 1
Gaps: 1
US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-564-357-17 (1-170)

QY 400 GGGCTGTGTGAAGAAAA--TCACCTTGTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTA 456

Db 1 G1yleuCy6glucllyysProAlaAlaLeuLeuProMeSerLeuSerGlnProCy6leu 20
QY 457 CCTGTGGCCAAACATTGGGCCCAACCCGAATTTCTCCCAATTTTATTTGCTGGCTCCAGCGA 516
Db 21 ProValProSerValG1yleuThrArg1leuProH1leuProH1leuGlySerGlnlys 40
QY 517 GATGTCTTCAACAAGAGCTGATATACAGCAAGATGGATGGTATATGTATGTTAAATGCCAGC 576
Db 41 AspValLeuAsnLysAspLeuMetThrGlnAsnGly1leSerTyValLeuAsnAlaSer 60
QY 577 TATACCTGTCCAAAGCCTGACTTATCCCGAGTCTCATTTCCGTGCTGCTGGAAT 636
Db 61 AsnSerCy6ProLysProAspPhe1leCy6GluSerArgPheMetArgValPro1leAsn 80
QY 637 GACAGCTTTGTGAGAAAAATTTGGCGTGGTGGCAAAATCAGATGATTGATTGAGAA 696
Db 81 AspAsnTyCy6gluLysLeuLeuProTyrLeuAspLysSer1leGluPhe1leAspLys 100
QY 697 GCAAAAGCCTCCAAATGATGTGTCTAGTCACTGTTAGCTGGATCTCCGCTCGCC 756
Db 101 AlAlyLeuSerSerCy6gluVal1leValH1sCy6leuAlaGly1leSerArgSerAla 120
QY 757 ACCATGCTATGCGCTACATCATGAGAGATGACATGCTTTAATGATGAAGCTTACAGA 816
Db 121 Thr1leAla1leAlaTyrlleMetLysThrMetGlyMeSerSerAspAlaTyArg 140
QY 817 TTTGTGAAGAAAAAGACCTACTATATCTCAAACTTCAATTTTCTGGGCCAACTCTG 876
Db 141 PheValLysAspArgArgProSer1leSerProAsnPheAsn1leuGlyGlnLeu 160
QY 877 GACTATGAGAAGATTAG 897
Db 161 GluTyrgluArgThrLeuLys 167

RESULT 5
US-09-619-380-16
; Sequence 16, Application US/09619380
; Patent No. 6643921
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; TITLE OF INVENTION: DSP-11 DUAL SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.418
; CURRENT APPLICATION NUMBER: US/09/619,380
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-619-380-16

Alignment Scores:
Pred. No.: 2,22e-71 Length: 170
Score: 654.50 Matches: 121
Percent Similarity: 88.62% Conserves: 27
Best Local Similarity: 72.46% Mismatches: 18
Query Match: 40.23% Indels: 1
Gaps: 1
US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-619-380-16 (1-170)

QY 400 GGGCTGTGTGAAGAAAA--TCACCTTGTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTA 456
Db 1 G1yleuCy6glucllyysProAlaAlaLeuLeuProMeSerLeuSerGlnProCy6leu 20
QY 457 CCTGTGGCCAAACATTGGGCCCAACCCGAATTTCTCCCAATTTTATTTGCTGGCTCCAGCGA 516
Db 21 ProValProSerValG1yleuThrArg1leuProH1leuProH1leuGlySerGlnlys 40
QY 517 GATGTCTTCAACAAGAGCTGATATACAGCAAGATGGATGGTATATGTATGTTAAATGCCAGC 576

Db 41 AsphAlIeunAenAynsAspIeunwEtThrGInAnGIlYIleSerTYrValIeunAsAlaSer 65
 Oy 577 TAIAcCTGTCcAAAGcCTGACTTAAATCCCGAGTCTCAATTCCTCGcGTGcCTGTGAAT 65
 Db 61 AsnSerCyAspIroLySproAspPhelIeCySgluSerAspPhenMetAspValProIleAn 80
 Oy 637 GACACCTTTTGAGAGAAATTTTGCCGTGGTTGACAAATCAGAGATTCATTGAGAA 69
 Db 81 AspAsenTYrCyGluLyluLeuNeuTroTYrIleuAspLySerIleGluPhelIeAspLyS 10
 Oy 697 GCAAAAGcCTCCAAAGATGTGTCTAGTGACCTGTTTACCTGGAGATCTCCGCTCCGC 75
 Db 101 AlAluYsLeuSerSerCyGlnValIleValHIeCySleuAlaIylIleSerTrSerAla 120
 Oy 757 ACCATCGCTATCGcCTCAATCATGAAAGAGATGACATGCTCTTATGATGAACCTTACGA 81
 Db 121 ThrIleAlaIleAlaTYrIleMelYsTrHMeGIlYMetSerSerAspAspAlaTYrArg 140
 Oy 817 TTGTGAAAGAAAAAGACCTACTATATCTCCAAATTCOAATTTCTGGGcCAACTCTG 87
 Db 141 PheAlaIlyAspArgTrpProSerIleSerProAsnPhenAsnPhelGluGlnLeu 160
 Oy 877 GACTATGAGAGAAAGATTAG 897
 Db 161 GluTYrGluArgThrIleuLyS 167

```

RESULT 6
US-09-557-921-2
; Sequence 2, Application US/09557921
; Patent No. 6551610
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.416
; CURRENT APPLICATION NUMBER: US/09/557,921
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 482
; TYPE: PRT
; ORGANSIM: Homo sapiens
US-09-557-921-2

```

Alignment Scores:	
Pred. No.:	2,156-48
Score:	469.00
Percent Similarity:	57.724
Best Local Similarity:	35.914
Query Match:	28.834
DB:	4
Length:	488
Matches:	107
Conservative:	65
Mismatches:	96
Indels:	30
Gaps:	7

79 CMC CM 2 mm x 2 mm

US-09-557-921-2 (1-482)

AAATGAAAGCCGCAATTTGTGGATTAACAATACATCCAACTTTGGAAACCAAT	136
	
Db	ValIleIleAspCysArgProPheMetGluTyrAsnLysSerHisIleGlnGlyAlaVal	137
	
Qy	AAATGCAACTCTCT---AAGCTTATGAAGCCGAAGGTTCGCAACAGCAAGGTGTAAAT	139
	
Db	HisIleAsnGlySerAlaAspLysIleSerAspArgArgLeuGlnGlnGlyIleIleThrVal	193
	
Qy	ACAAAGCTCAATCCAG---CATTCAGCGAAACAAATGAAGTTGACATTTGATGCACTGACAGAG	196
	
Db	LeuAspPheLeuSerCysArgGlnGlyLysAspSerPheLysAlaGllIlePheSerLysGln	213
	
Qy	GTTGTAAGTTTACGATCAAGAGCTCCCAAGAAGATGTGCTCTCTCTTCGACAGCTTTTCTC	253
	
Db	IleIleValTyrAspGlnAsnThrAsnGlnProSerIleArgValMetProSerGlnProLeu	233
	
Qy	ACTGTACTCTCGGTAAACTGGAAGAAGAGCTTCAACTCTGTTCACCGTGGCTTGACAGTGGG	313
	

```

Db      253 HistLevalLeuGIuSerLeuLysArgGluGluYlyGluProLeuValLeuLysGlyGly  ::||| |||||
OY      373 TTGGAGATCTCTCGTGGTTGTTCCCGCCGCTGTGAAGAAA-----  ::||| |||||
Db      273 LeuSerSerPheIlySGlnAsnHISgluAsnLeuCySAspAsnSerLeuGluIneGlnGlu  ::||| |||||
OY      418 -----TTCACCTGTCGCTCCCTACCTGCAATTCT  |||||
Db      293 CysArgGluValGlyGlyAlaSerAlaAlaSerLeuLeuProGluProLeu  |||||
OY      445 CAGCTTGCTTAACT---GTGGCAACATTGGGCCCAACCCGAATTCTTCCCATCTTAT  3
Db      312 ---ProThrThrProAspIleGluAsnAlaGluLeuThrProIleLeuProPheLeu  3
OY      502 CTTCGCGCCAGAGAGATGTCTCAACAGAGAGTATACAGCAAGAAATGGAGTTGGTAT  5
Db      331 LeuGIYAsnGluIneAspAlaGlnAspLeuAspThrMetGlnAspLeuAsnIleGlyTyr  3
OY      562 GTGTAAATGCCAGC-----TATACCTGTCCAAAGCCTGATTATATCCC  6
Db      351 ValIleAsnValThrThnHisLeuProLeuIynHisTyrGluYlySGlyLeuPhe-----  3
OY      607 GAGTTCATTTCCTCGTGTGCTGCGCTGTGAATACAGCTTTGTGTGAATAATTTGCGCTGG  6
Db      369 -----AsnTyrTlyAspGluProAlaThrAspSerAsnLysGlnAsnLeuAspGlnTyr  38
OY      667 TTGGCAACAATAGATATTCAATGAGAAAGCAAAAGCCCTCAATGATGTGTCTAGTG  72
Db      387 PheGluGluAlaPheGluPheIleGluGluAlaHISgluCyGlyLysGlyLeuLeuIle  40
OY      727 CACTGTTAGGTGGAGATCCGCCCTCCGCCACCATGCTATGCGCTACATCATGAAGAG  78
Db      407 HisCyGlnIleGlyValSerArgSerAlaThrIleValIleAlaTlyLeuMetLysHis  42
OY      787 ATGCAKATGTCTTAAATGAAGCTTACATATTGGAAAGAAAAGACCTACTATATTT  84
Db      427 ThrArgMetMetThrAspAlaTyrLysPheValYlySGlyAspGProIleIleSer  44
OY      847 CCAACTTCATATTCTGGGCCAACCCTCCGACATATGAGAAAGAAATTAAGAC  90
Db      447 ProAsnLeuAsnPheMetCylIneLeuLeuGluInePheGluGluAspLeuAsn  96

```

US-08-530-290-23
Sequence 23, Application US/08530290
Patent No. 5958721
GENERAL INFORMATION:
APPLICANT: Marshall, Christopher John
APPLICANT: Ashworth, Alan
APPLICANT: Hughes, David Anthony
TITLE OF INVENTION: Methods for Screening of Sub-
TITLE OF INVENTION: Therapeutic Activity and Ye
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,290
FILING DATE: 14-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB94/00694
FILING DATE: 31-MAR-1994

```
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9402573.1
FILING DATE: 10-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9307250.2
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Baetian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 084611-000000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-530-290-23

Alignment Scores:
Pred. No.: 2,73e-43      Length: 394
Score: 427.00           Matches: 101
Percent Similarity: 52.88% Conservative: 64
Best Local Similarity: 32.37% Mismatches: 116
Query Match: 26.24%      Indels: 31
DB: 2                   Gaps: 8

US-10-029-345a-108_COPY_538_1443 (1-906) x US-08-530-290-23 (1-394)
QY 45 GGTCGCTCTGCTGGAAGTGAACGGA-----AAAGT 77
DB 26 G1yG1yAlG1yG1y-SerG1ySerH1sg1yThrLeuG1yLeuProSerG1yG1yScy 45
QY 78 GCTGCTAATGATGAGCCGATTTGGTAATACATATCCACATTTGGAGCCAT 137
DB 45 eLeuLeuLeuAPCyArGPrProPheLeuAlah1SerAlaG1yTr1leuG1ySerVa 65
QY 138 TAATATCAACTGCTCCAGCTTATGAGCGAGGTTCACAGACAAAGTTAAATTC 197
DB 65 laenValaGcyAanThrlleValaArgaGrg--AlaG1yG1ySerValSerleuG1 84
QY 198 AGAGCTCATC-----CAGCATTCAGCGAAACATAGCTTGAATTCAGTCAAGAA 251
DB 84 uG1n1leLeuProAlaG1uG1uValaArgAlaArgLeuArgSerG1yLeuTr1SerAl 104
QY 252 GGTGAGTATTAAGATCAAGCTCCCAAGATGTTGCTCTCTCTTCAAGCTGTTTCT 311
DB 104 aVal1leValaTr1aSPoluarGserProaArgAlaG1uSerleuArgG1uSPserThrya 124
QY 312 CACTGTAATCTGAGTAACTGGAAGAGCTTC-----AACTGTTGCACTGCTTGC 365
DB 124 lSerleuValaValaG1a1a1euaArgaRgaenAlaG1uArgThraSP1leCySerleuG1y 144
QY 366 AGGTGGGTTTGTGAGTCTCTCTGTTGTTTCCCTGCGCTCTGGAAGAAATCACTCT 425
DB 144 sG1yG1yTr1G1uArgPheSerSerG1uTr1ProG1uPhCySerTr1Thrya1a1e 164
QY 426 A-----GTCCCTACCTGCAATTTCTCAGCCCT-----TGCTT 455
DB 164 uAla1a1a1leProProProAlaProProSerAlaThnG1uProLeuAPserleuAPCySe 184
QY 456 ACCTGTTGCC-----AACATTGGGCCCAACCCGAATTTCTTCCCATCTTTA 500
DB 184 rSerCyG1yThrProLeuH1aSPrg1nG1uG1yProValaG1uTr1leuProPheLeuTr1 204
QY 501 TCTTGGCTGCACGAGATCTCTCAACAAGAGCTGATACAGACGAATGGATTGGTTA 560
DB 204 r1euaG1ySerAlaTr1Th1a1a1a1aArgaSPMeleuAPAla1eudG1y1leuThra1 224
QY 561 TGTTTAAATGACGATATACCTGTCCAAAGCGCTGAATTCCTCCGAGTCAATTTCT 620
```

```
Db 224 aLeuLeuAenValaSerSerAPCyProAan---H1aPheG1uG1yH1eTr1G1nTr1y 243
QY 621 GCCTGTCCTGTGATGACAGCTTTGTGAGAAATTTTCCGTGGTGCACAAATCACT 680
DB 243 sCy1leProValaG1uAPaSPanH1aSP1a1aSP1leSerSerTr1PheMeG1uAla1 263
QY 681 AGATTTCATTTGAAGAAAGCAAAAGCCCTCAATGATGATGTTCTTCACTGTTAGCTGG 740
DB 263 eG1uTr1leAPa1aVala1yAPCyArG1yArgValaG1eValaH1eCySerG1nAlaG1 283
QY 741 GATCTCCGCTCCGCGACCATCCGCTATCGCTATCATATGAGAGATGACATGCTTT 800
DB 283 y1leSerArgSerAlaThrlleCySerleuAlaTr1leuMeuMetleYb1aArgValaArgle 303
QY 801 AGATGAAGCTTACAGATTTGTGAAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAG 860
DB 303 uG1uG1uAlaPhaG1uPhaVala1yG1nArgaSer1le1leSerProAanPheSerPh 323
QY 861 TCTGGGCCAATCTGTCGACATATGAGAAAGATT 894
DB 323 eWeG1yG1nLeuLeuG1nPhaG1uSerG1nVal 334

RESULT 8
US-09-702-705-805
; Sequence 805, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodges, Michael A.
; APPLICANT: Ranger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retler, Marc
; APPLICANT: Mammon, Jane
; APPLICANT: Pan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-705-805

Alignment Scores:
Pred. No.: 2,73e-43      Length: 394
Score: 427.00           Matches: 101
Percent Similarity: 52.88% Conservative: 64
Best Local Similarity: 32.37% Mismatches: 116
Query Match: 26.24%      Indels: 31
DB: 4                   Gaps: 8

US-10-029-345a-108_COPY_538_1443 (1-906) x US-09-702-705-805 (1-394)
QY 45 GGTCGCTCTGCTGGAAGTGAACGGA-----AAAGT 77
DB 26 G1yG1yAlG1yG1y-SerG1ySerH1sg1yThrLeuG1yLeuProSerG1yG1yScy 45
QY 78 GCTGCTAATGATGAGCCGATTTGGTAATACATATCCACATTTGGAGCCAT 137
DB 45 eLeuLeuLeuAPCyArGPrProPheLeuAlah1SerAlaG1yTr1leuG1ySerVa 65
QY 138 TAATATCAACTGCTCCAGCTTATGAGCGAGGTTCACAGACAAAGTTAAATTC 197
DB 65 laenValaGcyAanThrlleValaArgaGrg--AlaG1yG1ySerValSerleuG1 84
QY 198 AGAGCTCATC-----CAGCATTCAGCGAAACATAGTTGACATTCAGTCAAGAA 251
```

```

Db      84 ugnlileuProalaglunlvalargalalargleuargserglyleuYrSer1 104
QY      252 GGTGTGATTTAGACATCAAGCTCCAGAGATTTGCTCTCTCTCTTCAAGCTGTTTCT 311
Db      104 aValileVallyrAspIargserProargalagluserleuarglunAspserThrva 124
QY      312 CACTGACTTCTGGGTAACTGAGAGAGCTTC-----AAGCTGTTCACCTGCTTGC 365
Db      124 lserleuValIalglalalargAspAsnIagluArgThrAspIleCyseuLeuY 144
QY      366 AGGTGGTTTGTGAGTCTCTCTGTTTTCCTGGCTCTGTGAAGAAATTCACCT 425
Db      144 egllylyrIgluArgPheSerSerGluYrProglunPheCyseSerlyThryalale 164
QY      426 A-----GTCCCTACCTGACTTTCAGCCT-----TGCTT 455
Db      164 uAlalalleProProProValProProseralathrgIunPleuAspLeuAspCyse 184
QY      456 ACCTGTTGCC-----AACATGGGCGCAACCCGAATTCTTCCCAATCTTTA 500
Db      184 rSerCyseGlyThrProleuHIsaspIunlgluYrProvalglunleuProPheleuY 204
QY      501 TCTTGCTGCGCAGCAGATGTCCTCAACAGAGAGCTGATACAGCAATGGGATGGTTA 560
Db      204 rleuGlySerAlaYrHIsAlaAlaArgArgAspMetleuAspAlaleuGlyleThral 224
QY      561 TGTGTTAAATGCCAGCTATACCTGCAAGAGCTTATATCCCGAGCTCATTTTCC 620
Db      224 aleuLeuAsnValIserSerAspCyseProAsn--HIsPheglunIyHIsyrgIunYrly 243
QY      621 GCGTGGCTGTGAATGACAGCTTTTGTGAAGAAATTTTCCGCTGTGAGCAAACT 680
Db      243 sCyseIleProvalglunAspAsnHIsyAlaAspIleSerSerlyrPheMetGluAlaI 263
QY      681 AGATTTCAATTAGAAAGCAAAAGCTCAATGATGTTCTTGTGACCTGTTTACCTG 740
Db      263 eglunYrIleAspAlaValIyAspCyseArgGlyArgValleuValHIsCyseGluAlaI 283
QY      741 GATCCCGCTCGCCAGCAATGCTTATGCTTATCATCATGAAAGAGATGACATGCTT 800
Db      283 yIleSerIArgserAlaThrIleCyseleuAlaYrleuMetCetylsyAspValArgle 303
QY      801 AGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTATATCTCAAACTTCAATT 860
Db      303 uglunlAlaPheglunPheValIySglunArgIserIleIserProAsnPheSerPh 323
QY      861 TCTGGGCACTCTCTGAGCTATGAGAGAAAT 894
Db      323 eMetGlyGlnleuGlnPhegluserIunVal 334

RESULT 9
US-09-702-705-827
/ Sequence 827, Application US/09702705
/ Patent No. 6504010
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongrong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Panger, Gary
/ APPLICANT: Vedrick, Tom
/ APPLICANT: Carter, Darlick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C14
/ CURRENT APPLICATION NUMBER: US/09/702, 705
/ NUMBER OF SEQ ID NOS: 1833
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 827

```

```

/ LENGTH: 394
/ TYPE: PR1
/ ORGANISM: Homo sapiens
US-09-702-705-827

Alignment Scores:
Pred. No.: 2,73e-43
Score: 427.00
Percent Similarity: 53.85%
Best Local Similarity: 32.69%
Query Match: 26.24%
DB: 4
Gaps: 9

US-10-029-345a-108_COPY_538_1443 (1-906) x US-09-702-705-827 (1-394)
QY      45 GGTGCTCTGCTGCAAGATGGAAGCGAA-----AAAGT 77
Db      26 GlylylAlaGlyly-SerGlySerHIsGlylyrleuGlyleuProserGlylylyCy 45
QY      78 GCTCTAATGATGACCGGCAATTTGTGAATACATACATCCCATTTTGAAGCAAT 137
Db      45 sleuLeuAspCyseArgProPheleuAlaHIsSerAlaGlylyrIleleuGlyserVa 65
QY      138 TAATATCACTGCTCCAGCTTATGAGAGAGGTTGCAACAGCAAAAGTTAATTAC 197
Db      65 laenValArgCyseAsnThrIleValArgArg--AlaYseIySerValIserleuG 84
QY      198 AGAGCTCATC-----CAGCATTCAGCGAAACATPAGGTTGACATTGATTCAGTCAGAA 251
Db      84 ugnlileuProalaglunlglunlvalargalalargleuargserglyleuYrSer1 104
QY      252 GGTGTGATTTAGACATCAAGCTCCAGAGATTTGCTCTCTCTCTTCAAGCTGTTTCT 311
Db      104 aValileVallyrAspIargserProargalagluserleuarglunAspserThrva 124
QY      312 CACTGACTTCTGGGTAACTGAGAGAGCTTC-----AAGCTGTTCACCTGCTTGC 365
Db      124 lserleuValIalglalalargAspAsnIagluArgThrAspIleCyseuLeuY 144
QY      366 AGGTGGTTTGTGAGTCTCTCTGTTTTCCTGGCTCTGTGAAGAAATTCACCT 425
Db      144 egllylyrIgluArgPheSerSerGluYrProglunPheCyseSerlyThryalale 164
QY      426 A-----GTCCCTACCTGACTTTCAGCCT-----TGCTT 455
Db      164 uAlalalleProProProValProProseralathrgIunPleuAspLeuAspCyse 184
QY      451 ----TGC---TACCTGTTGCCAATTT---GGGCAACCGAATTTCTTCCCAATCTTTA 500
Db      184 rSerCyseGlyThrProleuHIsaspIunlgluYrProvalglunleuProPheleuY 204
QY      501 TCTTGCTGCGCAGCAGATGTCCTCAACAGAGAGCTGATACAGCAATGGGATGGTTA 560
Db      204 rleuGlySerAlaYrHIsAlaAlaArgArgAspMetleuAspAlaleuGlyleThral 224
QY      561 TGTGTTAAATGCCAGCTATACCTGCAAGAGCTTATATCCCGAGCTCATTTTCC 620
Db      224 aleuLeuAsnValIserSerAspCyseProAsn--HIsPheglunIyHIsyrgIunYrly 243
QY      621 GCGTGGCTGTGAATGACAGCTTTTGTGAAGAAATTTTCCGCTGTGAGCAAACT 680
Db      243 sCyseIleProvalglunAspAsnHIsyAlaAspIleSerSerlyrPheMetGluAlaI 263
QY      681 AGATTTCAATTAGAAAGCAAAAGCTTCAATGATGTTCTTGTGACCTGTTTACCTG 740
Db      263 eglunYrIleAspAlaValIyAspCyseArgGlyArgValleuValHIsCyseGluAlaI 283
QY      741 GATCCCGCTCGCCAGCAATGCTTATGCTTATCATCATGAAAGAGATGACATGCTT 800
Db      283 yIleSerIArgserAlaThrIleCyseleuAlaYrleuMetCetylsyAspValArgle 303
QY      801 AGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTATATCTCAAACTTCAATT 860

```

Db 303 ucgluglualaphnegluPhueVallysegluNargserlleleSerProanbheserPh 323
QY 861 TCTGGCCAACTCTGACATAGAGAAAGATT 894
Db 323 emecGlyInLeuInghnegluInsergluVal 334

RESULT 10
US-09-736-457-805
Sequence 805, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodges, Michael A.
APPLICANT: Ranger, Gary
APPLICANT: Vedrick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Pan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 805
LENGTH: 394
TYPE: PRN
ORGANISM: Homo sapiens
US-09-736-457-805

Alignment Scores:
Pred. No.: 2,73e-43 Length: 394
Score: 427.00 Matches: 101
Percent Similarity: 52.88% Conservative: 64
Best Local Similarity: 32.37% Mismatches: 116
Query Match: 26.24% Indels: 31
Gaps: 8
DB: 4

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-736-457-805 (1-394)

QY 45 GGTGGCTCTGCTGGAAGTGAACGAA-----AAAGT 77
Db 26 GtlyglYlaglYglY-SerGlYserHseglYthrlleuGlYleuProserGlYglYserCY 45
QY 78 GGTGTAATGATGAGCGCCATTGTGGAAATACATATCCACATTTTGGAGCCAT 137
Db 45 aleuLeuLeuAspCYarGpProPhleuAlahlsSerAlaglyTYrllleuGlYserVa 65
QY 138 TATATCAACTGCTCCAGAGTTTGAAGCGAGTTGCAACAGAGAAAGTTTAATAC 197
Db 65 lanValargCYaenThrllleValargarg---AlalYglYserValserleuGl 84
QY 198 AGAGCTCAGT-----CAGCATTCAGGAAACATAAGCTTGACATGATTCAGTAA 251
Db 84 ugnllleuProAlagluIngluValargAlaYglYleuargserGlYleuYrserAl 104
QY 252 GGTGTAGATTAGATCAAGATCCCAAGATGTGCTCTCTTCAAGACTGTTTTC 311
Db 104 aValllleValTYrAspGlYargserProargAlagluSerleuargluAspserThryVa 124
QY 312 CACTGACTTCTGGTAACTGGAAGAGCTTC-----AACTGTTCACCTGCTTC 365
Db 124 lserleuValValglAlaleuargargenAlagluargThAspilleCYseleuY 144
QY 366 AGGTGGTTGCTGAGATCTCTGTTGTTTCCCTGCGCTCTGGAAGAAATCCACTCT 425
Db 144 sglYglYTYrgluargPheserSerglYurYrProgluPhueCYserlyeThryAlale 164
QY 426 A-----GTCCCTACCTGATTTCTACAGCT-----TGCTT 455

Db 164 uAlaAlalleProProProValProProserAlathrlgluProleuAspLeuAspCyse 184
QY 456 ACCTGTGCC-----AACTGGGCCCAACCCGAATCTTCCCAATCTTTA 500
Db 184 rSerCYglYthrlProleuHlsarPglIngluIngluProValgluIlleleuProPhleuY 204
QY 501 TCTTGGCTGCAGCGAGATGTCTCAACAGAGCTGATACAGCAGATGGATGGTTA 560
Db 204 rleuGlYserAlaTYrHlsAlaAlaargargAspMetleuAspAlaleuGlYlthrlAl 224
QY 561 TGTGTAAATGCCAGSTATACTGTCCAAAGCTGATTTATCCCGAGCTCATTTCT 620
Db 224 aleuLeuAspValserSerAspCyserProan---HlsPhegluGlYHlsTYrGlInYrly 243
QY 621 GCCTGTCCCTGTGAATGACAGCTTTTGTGAGAAATTTTCCGTGTGACAAATCACT 660
Db 243 vCYsllleProValgluAspAspHlslyAlaAspIlleSerSerTrpPhmetolAla1 263
QY 681 AGATTCAATTGAGAAAGCAAAAGCCCTCAATGATGTGTTCTAGTCACTGTTAGCTGG 740
Db 263 egluTYrllleAspAlaVallyeAspCYargglYargValleuValHlsCYsegluAlagl 283
QY 741 GATTCCTCCGCTCGCCACCATGCTATCGCTACATCATGAGAGATGACATGCTTT 800
Db 283 YlleserArgserAlathrlleCYseuAlaTYrleuMetMetleYlYargValargle 303
QY 801 AGATGAGCTTACAGATTTGTGAAAGAAAAGACTTACTATTCCTCAACTTCATTT 860
Db 303 ucgluglualaphnegluPhueVallysegluNargserlleleSerProanbheserPh 323

QY 861 TCTGGCCAACTCTGACATAGAGAAAGATT 894
Db 323 emecGlyInLeuInghnegluInsergluVal 334

RESULT 11
US-09-736-457-827
Sequence 827, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodges, Michael A.
APPLICANT: Ranger, Gary
APPLICANT: Vedrick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Pan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 827
LENGTH: 394
TYPE: PRN
ORGANISM: Homo sapiens
US-09-736-457-827

Alignment Scores:
Pred. No.: 2,73e-43 Length: 394
Score: 427.00 Matches: 102
Percent Similarity: 53.85% Conservative: 66
Best Local Similarity: 32.69% Mismatches: 113
Query Match: 26.24% Indels: 31
Gaps: 9
DB: 4

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-736-457-827 (1-394)


```
QY 45 GGTGGCTCTGTGAAAGTGAACGAA-----AAAGT 77
DB 26 G1YGLYALAG1YGLY-SerGlySerHisGlyThrLeuGlyLeuProSerGlyYolYlyscY 45
QY 78 GGTGTAATGATTAAGCCGCAATTTGTGAATACATATACATATCCACATTTTGAAGCCAT 137
DB 45 sleuLeuLeuaspCysargProPheLeuAlaHisSerAlaGlyTrpIleLeuGlySerVa 65
QY 138 TAAATCAACTGCTCCAAAGCTTATGAAGCGAGTTGCAACAGACAAAGTGTATTAAC 197
DB 65 lAsnValArgCysasnThrIleValAlaGargArg--AlaYsGlySerValSerLeuG1 84
QY 198 AGAGCTATC-----CAGCATTCAGCGAAACATTAAGTTGACATTTGAGTGAAGTCAAA 251
DB 84 uGlnIleuProAlaGlnIleuGlnIleuAlaArgAlaArgLeuArgSerGlyLeuTrpSerAl 104
QY 252 GGTGTGATTTAGATCAATCAAGCTCCCAAGATGTGCTCTCTTCAAGCTTTTCT 311
DB 104 aValIleValIlyAspGlnArgSerProArgAlaGlnIleuLeuArgGlnAspSerThrVa 124
QY 312 CACGTACTTCTGGGTAACCTGGAAGAGCTTC-----AACTGTGTACCTGGTTC 365
DB 124 lSerIleuValAlaGlnAlaLeuArgAsnAlaGlnArgThrAspIleCysleuLeuY 144
QY 366 AGGTGGATTGTGAGATTTCTCGTTGTTCCCGCTCTGTGAAGAAATCCACTCT 425
DB 144 sGlyGlyTrpGluArgPheSerSerGlyTrpGlnPheCysSerIlyThrIlyAlaLe 164
QY 426 A-----GTCCCTACCTGATTTCTCAGCCT----- 450
DB 164 uAlaAlaIleProProProValProProSerAlaThrGlnProLeuAspLeuGlyCysSe 184
QY 451 ----TGC--TTACCTGTGTGCCACATTT--GGGCCAACCCGAATTTCTTCCAAATCTTTA 500
DB 184 rSerIySg1yThrProLeuHisAspGlnGlyYpProValGlnIleLeuProPheLeuY 204
QY 501 TCTTGCGCAGGAGGAGATGCTCCCAACAGAGCTGATACAGAGAAATGGAGATTGGTTA 560
DB 204 rLeuGlySerAlaTrpHisAlaAlaArgAspMetLeuAspAlaLeuGlyIleThrAl 224
QY 561 TGTGTAAATGCCAGCTATACCTGTCCCAAGCCTGATCTTATCCCGAGCTCATTTCT 620
DB 224 aLeuLeuAsnValSerSerAspSerProAsn---HisPheGlnGlyHisTrpGlnTrpYly 243
QY 621 GCGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTCCGTGTGTGACAATGAGT 680
DB 243 sCysIleProValGlnAspAsnHisIlysaAlaAspIleSerSerTrpPheMetGlnAlaI 263
QY 681 AGATTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTGTCTAGTGAACCTTTAGCTGG 740
DB 263 eGlnTrpIleAspAlaValIlyAspCysArgGlyArgValaLeuValHisCysGlnAlaG1 283
QY 741 GATCTCCGCTCCGCAACATGCTATGCGCTTACATCATGAAAGATGACATGTCTTT 800
DB 283 yIleSerArgSerAlaThrIleCysLeuAlaTrpLeuMetCmetIlysaArgValaArgLe 303
QY 801 AATAGAGCTTACAGATTGTGAAAGAAAGAAAGCTACTATATCTTCCAAATCTTCAATT 860
DB 303 uGlnGlnAlaPheGlnPheValIlySglnArgArgSerIleIleSerProAsnPhSerPh 323
QY 861 TCTGGGCCAATCTCCGAGCTATGAGAAAGAAAT 894
DB 323 eMetGlyGlnIleuLeuGlnPheGlnSerGlnVal 334
RESULT 12
US-09-614-124B-805
; Sequence 805, Application US/09614124B
; Parent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
```

```
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mamion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.478C9
CURRENT APPLICATION NUMBER: US/09/614,124B
CURRENT FILING DATE: 2001-07-11
NUMBER OF SEQ ID NOS: 1668
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 805
LENGTH: 394
TYPE: PRF
ORGANISM: Homo sapiens
US-09-614-124B-805
Alignment Scores:
Pred. No.: 2,73e-43
Score: 427.00
Percent Similarity: 52.68%
Best Local Similarity: 32.37%
Query Match: 26.24%
DB: 4
Gaps: 8
US-10-029-345a-108_copy_538_1443 (1-906) x US-09-614-124B-805 (1-394)
QY 45 GGTGGCTCTGTGAAAGTGAACGAA-----AAAGT 77
DB 26 G1YGLYALAG1YGLY-SerGlySerHisGlyThrLeuGlyLeuProSerGlyYolYlyscY 45
QY 78 GGTGTAATGATTAAGCCGCAATTTGTGAATACATATACATATCCACATTTTGAAGCCAT 137
DB 45 sleuLeuLeuaspCysargProPheLeuAlaHisSerAlaGlyTrpIleLeuGlySerVa 65
QY 138 TAAATCAACTGCTCCAAAGCTTATGAAGCGAGTTGCAACAGACAAAGTGTATTAAC 197
DB 65 lAsnValArgCysasnThrIleValAlaGargArg--AlaYsGlySerValSerLeuG1 84
QY 198 AGAGCTATC-----CAGCATTCAGCGAAACATTAAGTTGACATTTGAGTGAAGTCAAA 251
DB 84 uGlnIleuProAlaGlnIleuGlnIleuAlaArgAlaArgLeuArgSerGlyLeuTrpSerAl 104
QY 252 GGTGTGATTTAGATCAATCAAGCTCCCAAGATGTGCTCTCTTCAAGCTTTTCT 311
DB 104 aValIleValIlyAspGlnArgSerProArgAlaGlnIleuLeuArgGlnAspSerThrVa 124
QY 312 CACGTACTTCTGGGTAACCTGGAAGAGCTTC-----AACTGTGTACCTGGTTC 365
DB 124 lSerIleuValAlaGlnAlaLeuArgAsnAlaGlnArgThrAspIleCysleuLeuY 144
QY 366 AGGTGGATTGTGAGATTTCTCGTTGTTCCCGCTCTGTGAAGAAATCCACTCT 425
DB 144 sGlyGlyTrpGluArgPheSerSerGlyTrpGlnPheCysSerIlyThrIlyAlaLe 164
QY 426 A-----GTCCCTACCTGATTTCTCAGCCT----- 450
DB 164 uAlaAlaIleProProProValProProSerAlaThrGlnProLeuAspLeuGlyCysSe 184
QY 451 ----TGC--TTACCTGTGTGCCACATTT--GGGCCAACCCGAATTTCTTCCAAATCTTTA 500
DB 184 rSerIySg1yThrProLeuHisAspGlnGlyYpProValGlnIleLeuProPheLeuY 204
QY 501 TCTTGCGCAGGAGGAGATGCTCCCAACAGAGCTGATACAGAGAAATGGAGATTGGTTA 560
DB 204 rLeuGlySerAlaTrpHisAlaAlaArgAspMetLeuAspAlaLeuGlyIleThrAl 224
QY 561 TGTGTAAATGCCAGCTATACCTGTCCCAAGCCTGATCTTATCCCGAGCTCATTTCT 620
DB 224 aLeuLeuAsnValSerSerAspSerProAsn---HisPheGlnGlyHisTrpGlnTrpYly 243
QY 621 GCGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTCCGTGTGTGACAATGAGT 680
DB 243 sCysIleProValGlnAspAsnHisIlysaAlaAspIleSerSerTrpPheMetGlnAlaI 263
QY 681 AGATTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTGTCTAGTGAACCTTTAGCTGG 740
DB 263 eGlnTrpIleAspAlaValIlyAspCysArgGlyArgValaLeuValHisCysGlnAlaG1 283
QY 741 GATCTCCGCTCCGCAACATGCTATGCGCTTACATCATGAAAGATGACATGTCTTT 800
DB 283 yIleSerArgSerAlaThrIleCysLeuAlaTrpLeuMetCmetIlysaArgValaArgLe 303
QY 801 AATAGAGCTTACAGATTGTGAAAGAAAGAAAGCTACTATATCTTCCAAATCTTCAATT 860
DB 303 uGlnGlnAlaPheGlnPheValIlySglnArgArgSerIleIleSerProAsnPhSerPh 323
QY 861 TCTGGGCCAATCTCCGAGCTATGAGAAAGAAAT 894
DB 323 eMetGlyGlnIleuLeuGlnPheGlnSerGlnVal 334
```

```
Db 243 bcysileProvalGluAspAraHnHslybAlaSerIleSerSerTrpPheMetGluAlaI 263
Qy 681 AGATTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTGTTTCTGACTGTTTACTGG 740
Db 263 egutyrIleAraPAlaValIlyAspCybaArgIyArGValLeuValHsCyseGlnIaG 283
Qy 741 GATCTCCGCTCGCCAGCAGCATGCTATCGCTTACATCATCAAGAGATGACATGCTTT 800
Db 283 yIleSerAraSerIleThrIleCyseuAlaTrpLeuMetClybLysArgValArGle 303
Qy 801 AGATGAAGCTTACAGATTGTGTGAAGAAAAGACCTACTATATCTCCAACTTCATT 860
Db 303 uGlnGluAlaPheGluPheGluAlaValyGlnArGArGSerIleIleSerProAspHsSerPh 323
Qy 861 TCTGGGCGCAACTCTTGACTATGAGAAGATT 894
Db 323 emetGlyGlnLeuGlnPheGlnSerGlnVal 334

RESULT 13
US-09-614-124B-827
/ Sequence 827, Application US/09614124B
/ Patent No. 6630574
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mamon, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ FILE REFERENCE: 210121.478C9
/ CURRENT APPLICATION NUMBER: US/09/614,124B
/ CURRENT FILING DATE: 2001-07-11
/ NUMBER OF SEQ ID NOS: 1668
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 827
/ LENGTH: 394
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-614-124B-827

Alignment Scores:
Pred. No.: 2,73e-43 Length: 394
Score: 427.00 Matches: 102
Percent Similarity: 53.85% Conservative: 66
Best Local Similarity: 32.69% Mismatches: 113
Query Match: 26.24% Indels: 31
DB: 4 Gaps: 9

US-10-029-345A-108_COPY_538_1443 (1-906) x US-09-614-124B-827 (1-394)
Qy 45 GGTGGCTGCTGGAGAAAGTGAACGAA-----AAAGT 77
Db 26 GlyGlyAlaGlyGly-SerGlySerHsGlyThrLeuGlyLeuProSerGlyGlyLysCy 45
Qy 78 GGTGCTAATGTGATAGCCGCGCATTTGTGAAATACATATCCACATTTTGGAGCCAT 137
Db 45 aleuLeuLeuAraPcybaArgProheLeuAlaHsSerAlaGlyTrpIleLeuGlySerVa 65
Qy 138 TAATATCAACTGCTCCAGAGCTTATGAAAGGAGAGTTCGAACAGACAAAGGTAAATTC 197
Db 65 lAenValArGcybaArHsThrIleValArGArGArg--AlaLyGlySerValSerLeuG 84
Qy 198 AGAGCTCAGT-----CAGCATTCAAGGAAACATAGTTCAGATTGATTCAGTCAAG 251
Db 84 uGlnIleLeuProAlaGlnGluAlaValArGAlaArGLeuArGSerGlyLeuTrpSerAl 104
Qy 252 GGTGTAGATTAGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCAGACTGTTTCT 311
Db 104 aValIleValIlyArAspGluArGSerProArGAlaGlnSerLeuArGLeuAspSerThrVa 124
```

```
Qy 312 CACTGACTCTTGCGGAAACGTGAGAGAGCTTC-----AACTGTTCACTGCTGCG 365
Db 124 lSerLeuValValGlnAlaLeuArGArGAsnAlaGluArGThrArpIleCyseLeuLeu 144
Qy 366 AGGTGGCTTGTGCTGATGCTCTGTTGTTCCCTGGCCCTGTGTGAAGAAAATCCACTCT 425
Db 144 GgIyGlyTrpGluArGPhSerSerGlyTrpProGluPheCyseLysThrValAla 164
Qy 426 A-----GTCCCTACTGCAATTCTCAAGCT----- 450
Db 164 uAlaAlaIleProProProProValProProSerAlaThrGluPLeuAspLeuGlyCyse 184
Qy 451 -----TTCCTGTTGCCAACAATT---GGGCGAACCCGAATTCCTCCCAATCTTA 500
Db 184 rSerGlyGlyThrProLeuHsArpGlnGlyLeuProValGluIleLeuProPheLeu 204
Qy 501 TCTTGAGCTGCGCAGAGATGTCCTCAACAGAGCTGATACAGCAATGAGATTGGTTA 560
Db 204 rLeuGlySerAlaTrpHsAlaAlaArGArGArpMetLeuArAlaLeuGlyIleThrAl 224
Qy 561 TGTGTTAAATGCCAGCTATACCTGTCCAAAGCCTGATTATCCCGAGTTCATTCTCT 620
Db 224 aleuLeuAraValSerSerArPcybaProAsn--HisPheGluGlyHsIleTrpGlnTrp 243
Qy 621 GCGTGGCCTGTGAATGACAGCTTTGTGAGAAAATTTGCGGTGTGAGCAAAATCAG 680
Db 243 bcysileProvalGluAspAraHnHslybAlaSerIleSerSerTrpPheMetGluAlaI 263
Qy 681 AGATTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTGTTTCTGACTGTTTACTGG 740
Db 263 egutyrIleAraPAlaValIlyAspCybaArgIyArGValLeuValHsCyseGlnIaG 283
Qy 741 GATCTCCGCTCGCCAGCAGCATGCTATCGCTTACATCATCAAGAGATGACATGCTTT 800
Db 283 yIleSerAraSerAlaThrIleCyseuAlaTrpLeuMetClybLysArgValArGle 303
Qy 801 AATGAAGCTTACAGATTGTGTGAAGAAAAGACCTACTATATCTCCAACTTCATT 860
Db 303 uGlnGluAlaPheGluPheGluAlaValyGlnArGArGSerIleIleSerProAspHsSerPh 323
Qy 861 TCTGGGCGCAACTCTTGACTATGAGAAGATT 894
Db 323 emetGlyGlnLeuGlnPheGlnSerGlnVal 334

RESULT 14
US-09-671-325-805
/ Sequence 805, Application US/09671325
/ Patent No. 6667154
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mamon, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C12
/ CURRENT APPLICATION NUMBER: US/09/671,325
/ CURRENT FILING DATE: 2000-09-26
/ NUMBER OF SEQ ID NOS: 1825
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 805
/ LENGTH: 394
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-671-325-805

Alignment Scores:
```



```
QY 501 TCTTGCTGCAGCAGATGTCTCTCAACAGAGCTGATACAGACAGATGGATTGTTA 560
Db 204 rleuglyserlatyHialalaaargahpmetleuaspalaleuGlylethrAl 224
QY 561 TGTGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTTATCCCGAGTCTCATTTCT 620
Db 224 aleuleuanvaliserSerharpCysProasn--HispheluglyHisTyrgIntyrlY 243
QY 621 GCCTGTCCTGTGAAATGACAGCTTTTGTGAGAAATTTGCCGTGGTTGACAAATCAGT 680
Db 243 sCysrileProvalGluaspasnHislybalaasrileSerSertrpPheMetGluAlaI 263
QY 681 AGATTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTGTTCTAGTCACTGTTAGCTG 740
Db 263 eglutyrileaspalavalyleasrPcyasrgilyargvalleuvalHisCysglAlaI 283
QY 741 GATCTCCGCTCCGCCACCATCGCTATCGCTACATCATGAAAGAGATGAGCATGTCTTT 800
Db 283 ylieserArgserAlaThrileCysleuAlaTyrlleuMetMetlyslasrgValArgle 303
QY 801 AGATGAAGCTTACAGATTGTGTGAAGAAAAAGACCTACTATATCTCCAAACTTCAATT 860
Db 303 uGluGluAlaPheGluPheValylsGlnArgSerIleIleSerProasnpheserPh 323
QY 861 TCTGGGCAACTCTGTGACTATGAGAAGAATTT 894
Db 323 emetGlyGlnleuGlnPheGlnSerGlnVal 334
```

Search completed: June 21, 2004, 12:43:28
Job time : 17.5553 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 12:17:10 ; Search time 5703.02 Seconds

(without alignment)
6885.611 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 906
Sequence: 1 atgagccatagatgatggtg.....agaagataagaccagact 906

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing filter 45 summaries

Database :
1: gb ba:*
2: gb hcg:*
3: gb in:*
4: gb ov:*
5: gb ov:*
6: gb pac:*
7: gb ph:*
8: gb pl:*
9: gb pr:*
10: gb pr:*
11: gb ste:*
12: gb sy:*
13: gb un:*
14: gb vi:*
15: gb ba:*
16: em fun:*
17: em fun:*
18: em in:*
19: em mu:*
20: em om:*
21: em or:*
22: em ov:*
23: em pac:*
24: em ph:*
25: em pl:*
26: em ro:*
27: em ste:*
28: em un:*
29: em vi:*
30: em hcg hum:*
31: em hcg hum:*
32: em hcg other:*
33: em hcg mus:*
34: em hcg pin:*
35: em hcg rod:*
36: em hcg mam:*
37: em hcg vrt:*
38: em sy:*
39: em hcg hum:*
40: em hcg mus:*
41: em hcg other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	906	100.0	5450	6	AX482439
2	906	100.0	5450	6	AX482478
3	902.8	99.6	1988	6	AX260342
4	902.8	99.6	2071	6	AX921917
5	902.8	99.6	2118	6	AX099933
6	902.8	99.6	2732	6	AX180875
7	902.8	99.6	3059	6	AX278461
8	902.8	99.6	3284	6	BC042101
9	902.8	99.6	3496	6	AX441210
10	902.8	99.6	3521	6	AB052156
11	902.8	99.6	3544	6	AX260340
12	902.8	99.6	3566	6	AF506796
13	902.8	99.6	3766	6	AX374994
14	902.8	99.6	4790	6	BD171157
15	902.8	99.6	4790	6	BD183422
16	902.8	99.6	4790	6	AB051487
17	901.2	99.5	2102	6	AX713989
18	901.2	99.5	2102	9	AK055973
19	901.2	99.5	3104	6	AX405700
20	901.2	99.5	5111	6	AX482372
21	817.8	90.3	2200	6	AX921915
22	797.2	88.0	4874	10	AB052157
23	797.2	88.0	4943	10	BC059232
24	797.2	88.0	4975	10	BC057321
25	794	87.6	4992	10	AF345951
26	782.8	86.4	2756	6	AX482444
27	756.8	83.5	2807	6	BC031643
28	727.6	80.3	749	6	AX868469
29	727.6	80.3	749	6	BD148531
30	714.2	78.8	4026	10	AF345953
31	564.8	62.3	1935	9	AY038927
32	564.8	62.3	3332	6	AX441229
33	463.6	51.2	4837	10	AF345952
34	404.8	44.7	3231	5	BC063941
35	383.8	42.4	3661	10	AF345954
36	365.8	40.4	2377	6	AR255944
37	365.8	40.4	2377	9	HS027193
38	364.2	40.2	2264	6	BC045110
39	363.6	40.1	2453	6	A59888
40	363.6	40.1	2453	10	MMNTTPIGN
41	363.6	40.1	2600	10	BC052705
42	351	38.7	2415	6	A59887
43	335.8	37.1	2476	6	AX552119
44	317	35.0	408	6	AX482446
45	311.6	34.4	183797	9	AC022400

ALIGNMENTS

RESULT 1	LOCUS	AX482439	5450 bp	DNA	linear	PAT 16-AUG-2002
AX482439	DEFINITION	Sequence 108 from Patent WO02057460.				
AX482439	ACCESSION	AX482439				
AX482439.1	VERSION	GI:22316984				
KEYWORDS	SOURCE					
ORGANISM	ORGANISM	Homo sapiens (human)				
REFERENCE	REFERENCE					
AUTHORS	AUTHORS	Toderud,C.G., Bol,D., Finger,J., Lee,J., Nelson,T., Schieven,G., Suchard,S., Banae,D., Bassolino,D., Feder,J., Krystek,S., Mcatee,P., Mintler,G., Siemers,N., Jackson,D.G. and Ramanathan,C.				

Pred. No. is the number of results predicted by chance to have a

538 ATGGCCCATGAGATGATGGAACTCAAAATTGTTACTGAGAGGTTGGGCTCTGCTGAA 597
QY 61 AGTGAACGGAAGAGTCTGCTAATTGATGACCGGCATTGTGGAAATCAATACATCC 120
DB 598 AGTGAACGGAAGAGTCTGCTAATTGATGACCGGCATTGTGGAAATCAATACATCC 657
QY 121 CACATTTTGGAAACCAATTAATACACTGCTCAAGCTTATGAAAGGAGTTGCAACG 180
DB 658 CACATTTTGGAAACCAATTAATACACTGCTCAAGCTTATGAAAGGAGTTGCAACG 717
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAATAAGTTGACATGAT 240
DB 718 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAATAAGTTGACATGAT 777
QY 241 TGCAGTCAGAGGTTGTATGATTAAGATCAAAAGTCCCAAGATGTTGCTCTCTTCA 300
DB 778 TGCAGTCAGAGGTTGTATGATTAAGATCAAAAGTCCCAAGATGTTGCTCTCTTCA 837
QY 301 GACTGTTTCTCACTGCTACTCTGGGTTAACTGGAAGAGCTTCAACTCTGTTCACTG 360
DB 838 GACTGTTTCTCACTGCTACTCTGGGTTAACTGGAAGAGCTTCAACTCTGTTCACTG 897
QY 361 CTTCAGAGTGGTGTCTGAGTTCCTCTGTTTCCCTGCTCTGTAAGGAAATCC 420
DB 898 CTTCAGAGTGGTGTCTGAGTTCCTCTGTTTCCCTGCTCTGTAAGGAAATCC 957
QY 421 ACTTAGTCCCTAAGCTGATTTCTCAGCTTGTCTTACTGTTGCCAATTTGGGCCAAC 480
DB 958 ACTTAGTCCCTAAGCTGATTTCTCAGCTTGTCTTACTGTTGCCAATTTGGGCCAAC 1017
QY 481 CGAATTTCTCCCAATCTTATCTTGGCTGCGACGAGATGCTCTCAAGAGGCTGATA 540
DB 1018 CGAATTTCTCCCAATCTTATCTTGGCTGCGACGAGATGCTCTCAAGAGGCTGATA 1077
QY 541 CAGCAGAAATGGAGTGGTATGTTAATGACAGCTATCTGCTCAAGAGCTGACTT 600
DB 1078 CAGCAGAAATGGAGTGGTATGTTAATGACAGCTATCTGCTCAAGAGCTGACTT 1137
QY 601 ATCCCGAGTCTCAATTTCTGCGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 660
DB 1138 ATCCCGAGTCTCAATTTCTGCGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 1197
QY 661 CCGTGGTGGACAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCTCCAAATGATGTT 720
DB 1198 CCGTGGTGGACAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCTCCAAATGATGTT 1257
QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCAACATCCCTATCCCTATCATCATG 780
DB 1258 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCAACATCCCTATCCCTATCATCATG 1317
QY 781 AAGAGATGACATGTTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
DB 1318 AAGAGATGACATGTTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1377
QY 841 ATATCTCCAAATCTTCAATTTTCTGGGCAATCTCTGACATGAGAAAGATTAAGAAC 900
DB 1378 ATATCTCCAAATCTTCAATTTTCTGGGCAATCTCTGACATGAGAAAGATTAAGAAC 1437
QY 901 CAGACT 906
DB 1438 CAGACT 1443

RESULT 3
AX260342 1998 bp DNA linear PAT 26-OCT-2001
LOCUS Sequence 3 from Patent WO0173059.
DEFINITION AX260342
ACCESSION AX260342.1 GI:16509305
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS Meyers, R.A.
TITLE 3862 and 21117: dual specificity phosphatase molecules and uses
therefor
JOURNAL Patent: WO 0173059-A 3 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifier
1..1998
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.6%; Score 902.8; DB 6; Length 1998;
Best Local Similarity 99.8%; Pred. No. 4, 6e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAACTCAAAATTGTTACTGAGAGGTTGGGCTCTGCTGAA 60
DB 1 ATGGCCCATGAGATGATGGAACTCAAAATTGTTACTGAGAGGTTGGGCTCTGCTGAA 60
QY 61 AGTGAACGGAAGAGTCTGCTAATTGATGACCGGCATTGTGGAAATCAATACATCC 120
DB 61 AGTGAACGGAAGAGTCTGCTAATTGATGACCGGCATTGTGGAAATCAATACATCC 120
QY 121 CACATTTTGGAAACCAATTAATACACTGCTCAAGCTTATGAAAGGAGTTGCAACG 180
DB 121 CACATTTTGGAAACCAATTAATACACTGCTCAAGCTTATGAAAGGAGTTGCAACG 180
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAATAAGTTGACATGAT 240
DB 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAATAAGTTGACATGAT 240
QY 241 TGCAGTCAGAGGTTGTATGATTAAGATCAAAAGTCCCAAGATGTTGCTCTCTTCA 300
DB 241 TGCAGTCAGAGGTTGTATGATTAAGATCAAAAGTCCCAAGATGTTGCTCTCTTCA 300
QY 301 GACTGTTTCTCACTGCTACTCTGGGTTAACTGGAAGAGCTTCAACTCTGTTCACTG 360
DB 301 GACTGTTTCTCACTGCTACTCTGGGTTAACTGGAAGAGCTTCAACTCTGTTCACTG 360
QY 361 CTTCAGAGTGGTGTCTGAGTTCCTCTGTTTCCCTGCTCTGTAAGGAAATCC 420
DB 361 CTTCAGAGTGGTGTCTGAGTTCCTCTGTTTCCCTGCTCTGTAAGGAAATCC 420
QY 421 ACTTAGTCCCTAAGCTGATTTCTCAGCTTGTCTTACTGTTGCCAATTTGGGCCAAC 480
DB 421 ACTTAGTCCCTAAGCTGATTTCTCAGCTTGTCTTACTGTTGCCAATTTGGGCCAAC 480
QY 481 CGAATTTCTCCCAATCTTATCTTGGCTGCGACGAGATGCTCTCAAGAGGCTGATA 540
DB 481 CGAATTTCTCCCAATCTTATCTTGGCTGCGACGAGATGCTCTCAAGAGGCTGATA 540
QY 541 CAGCAGAAATGGAGTGGTATGTTAATGACAGTATACCTGCAAAAGCTTACT 600
DB 541 CAGCAGAAATGGAGTGGTATGTTAATGACAGTATACCTGCAAAAGCTTACT 600
QY 601 ATCCCGAGTCTCAATTTCTGCGCTGCTGCTGATGACAGCTTTTGTGAAGAAATTTTG 660
DB 601 ATCCCGAGTCTCAATTTCTGCGCTGCTGCTGATGACAGCTTTTGTGAAGAAATTTTG 660
QY 661 CCGTGGTGGACAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCTCCAAATGATGTT 720
DB 661 CCGTGGTGGACAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCTCCAAATGATGTT 720
QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCAACATCCCTATCATCATG 780
DB 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCAACATCCCTATCATCATG 780
QY 781 AAGAGATGACATGTTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
DB 781 AAGAGATGACATGTTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840

QY 841 ATATCTCCAACTTCATTTTCTGGCCAACTCTGAGCTATGAGAAGATTAAAGAC 900
 Db 841 ATATCTCCAACTTCATTTTCTGGCCAACTCTGAGCTATGAGAAGATTAAAGAC 900
 QY 901 CAGACT 906
 Db 901 CAGACT 906

RESULT 4
 AX921917 2071 bp DNA linear PAT 18-DEC-2003
 LOCUS Sequence 257 from Patent WO02068649.
 DEFINITION AX921917
 ACCESSION AX921917 GI:40215410
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 Patent: WO 02068649-A 257 06-SEP-2002;
 AUTHORS Curegen Corporation (US)
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1.2071
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 99.6%; Score 902.8; DB 6; Length 2071;
 Best Local Similarity 99.8%; Pred. No. 4.6e-267;
 Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGAACTCAATTTGTTACTGAGGTTGGCTCTGCTGGA 60
 Db 61 ATGGCCCATGAGATGATGAACTCAATTTGTTACTGAGGTTGGCTCTGCTGGA 60
 QY 61 ATGGGAACGAAAAGTCTGCTTAATGATGAGCCGCAATTTGGGATTAACATTC 120
 Db 121 AGTGAACGAAAAGTCTGCTTAATGATGAGCCGCAATTTGGGATTAACATTC 120
 QY 121 CACATTTTGAAGCCATTATCACTGCTCCAGCTTATGACGGAAGTTGCAATCC 180
 Db 121 CACATTTTGAAGCCATTATCACTGCTCCAGCTTATGACGGAAGTTGCAATCC 180
 QY 181 CACATTTTGAAGCCATTATCACTGCTCCAGCTTATGACGGAAGTTGCAATCC 240
 Db 181 CACATTTTGAAGCCATTATCACTGCTCCAGCTTATGACGGAAGTTGCAATCC 240
 QY 241 TGCAGTCAAGAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 300
 Db 241 TGCAGTCAAGAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 301 TGCAGTCAAGAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 360
 Db 301 TGCAGTCAAGAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 361 GACTGTTTCTCACTGTTCTCTGTTGTTTCTGTTGTTTCTGTTGTTTCTGTT 420
 Db 361 GACTGTTTCTCACTGTTCTCTGTTGTTTCTGTTGTTTCTGTTGTTTCTGTT 420
 QY 421 CTGTCAGTGGGTTTCTGTTGTTTCTGTTGTTTCTGTTGTTTCTGTTGTTTCT 480
 Db 421 CTGTCAGTGGGTTTCTGTTGTTTCTGTTGTTTCTGTTGTTTCTGTTGTTTCT 480
 QY 481 ACTCTAGTCCCTACCTGCTATTTCTGAGCTTCTGTTGCTTCTGTTGCTTCT 540
 Db 481 ACTCTAGTCCCTACCTGCTATTTCTGAGCTTCTGTTGCTTCTGTTGCTTCT 540
 QY 541 CGAATTTCTCCCAATCTTTTCTGTTGCTTCTGTTGCTTCTGTTGCTTCTGTT 600
 Db 541 CGAATTTCTCCCAATCTTTTCTGTTGCTTCTGTTGCTTCTGTTGCTTCTGTT 600
 QY 600 GACGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 Db 600 GACGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

Db 601 CAGCAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 601 ATCCCGAGTCTCATTTTCTGAGTGGTCTGTTGAATGACACCTTTTGGAAAAATTTG 660
 Db 661 ATCCCGAGTCTCATTTTCTGAGTGGTCTGTTGAATGACACCTTTTGGAAAAATTTG 720
 QY 720 CCGTGGTGGCAAAATCAGTATGATTTCAATTGAGAAACAAAGGCTCCATGATGTT 720
 Db 721 CCGTGGTGGCAAAATCAGTATGATTTCAATTGAGAAACAAAGGCTCCATGATGTT 780
 QY 780 CTAAGTCACTGTTTATGCTGAGTCTCCGCTCCGCAACCAATGCTTATGATGATG 840
 Db 781 CTAAGTCACTGTTTATGCTGAGTCTCCGCTCCGCAACCAATGCTTATGATGATG 840
 QY 840 AAGAGATGACATGCTTTTATGATGAAAGCTTACAGATTTTGTGAAAAGAAAAGCTTACT 840
 Db 841 AAGAGATGACATGCTTTTATGATGAAAGCTTACAGATTTTGTGAAAAGAAAAGCTTACT 900
 QY 900 ATATCTCCAACTTCATTTTCTGGCCAACTCTGAGCTATGAGAAGATTAAAGAC 900
 Db 901 ATATCTCCAACTTCATTTTCTGGCCAACTCTGAGCTATGAGAAGATTAAAGAC 960
 QY 901 CAGACT 906
 Db 961 CAGACT 966

RESULT 5
 AX099933 2118 bp DNA linear PAT 02-APR-2001
 LOCUS Sequence 15 from Patent WO0120004.
 DEFINITION AX099933
 ACCESSION AX099933
 VERSION AX099933.1 GI:13538943
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 Yue, H., Tang, Y. T., Bandman, O., Hillman, D. L., Baughn, M. R.,
 Azimzal, Y., and Lu, D. A.
 TITLE Protein phosphatase and kinase proteins
 JOURNAL Patent: WO 0120004-A 15 22-MAR-2001;
 INCYTE Genomics, Inc. (US)
 FEATURES
 source
 1.2118
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID NO: 1234795CB1"

ORIGIN

Query Match 99.6%; Score 902.8; DB 6; Length 2118;
 Best Local Similarity 99.8%; Pred. No. 4.6e-267;
 Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGAACTCAATTTGTTACTGAGGTTGGCTCTGCTGGA 60
 Db 25 ATGGCCCATGAGATGATGAACTCAATTTGTTACTGAGGTTGGCTCTGCTGGA 60
 QY 61 ATGGGAACGAAAAGTCTGCTTAATGATGAGCCGCAATTTGGGATTAACATTC 120
 Db 85 ATGGGAACGAAAAGTCTGCTTAATGATGAGCCGCAATTTGGGATTAACATTC 144
 QY 144 CACATTTTGAAGCCATTATCACTGCTCCAGCTTATGACGGAAGTTGCAATCC 180
 Db 145 CACATTTTGAAGCCATTATCACTGCTCCAGCTTATGACGGAAGTTGCAATCC 204
 QY 204 GACAAAGTTTAAATCAAGAGCTTCAAGATTCAGAGAAACATTAAGTTGATGAT 240
 Db 205 GACAAAGTTTAAATCAAGAGCTTCAAGATTCAGAGAAACATTAAGTTGATGAT 264

QY 241 TGCAGTGAAGAGTGTAGTTTACATCAAGCTCCAGATGTTGCTCTCTTCA 300
DB 265 TGCAGTGAAGAGTGTAGTTTACATCAAGCTCCAGATGTTGCTCTCTTCA 324
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAGAGCTTCACTCTGACCTG 360
DB 325 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAGAGCTTCACTCTGACCTG 384
QY 361 CTTCAGAGTGGGTTTCTGATGTTCTCTGTTTCTTCCCTGCTCTGTAAGAAATTC 420
DB 385 CTTCAGAGTGGGTTTCTGATGTTCTCTGTTTCTTCCCTGCTCTGTAAGAAATTC 444
QY 421 ACTCTAGTCCCTACCTGATCTTCTGACCTTGTACCTGCTGCAATTTGGGCAACC 480
DB 445 ACTCTAGTCCCTACCTGATCTTCTGACCTTGTACCTGCTGCAATTTGGGCAACC 504
QY 481 CGAATCTTCCCACTCTTATCTTGGCTGCGAGAGATGCTCAACAAGAGCTGATA 540
DB 505 CGAATCTTCCCACTCTTATCTTGGCTGCGAGAGATGCTCAACAAGAGCTGATA 564
QY 541 CAGCAGATGGGATGGTTATGTTTAAATGCGACGTATACCTGTCCAAAGCTGACTT 600
DB 565 CAGCAGATGGGATGGTTATGTTTAAATGCGACGTATACCTGTCCAAAGCTGACTT 624
QY 601 ATCCCGAGTCTCATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 660
DB 625 ATCCCGAGTCTCATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 684
QY 661 CCGTGTGTGACAAATCATGATGATTTCTGAGAAAGCAAAAGCTTCAATGATGTT 720
DB 685 CCGTGTGTGACAAATCATGATGATTTCTGAGAAAGCAAAAGCTTCAATGATGTT 744
QY 721 CTAGTCACTGTTTATGCTGGGATCTCCGCTCCGCAACCTGATATGCTTACATATG 780
DB 745 CTAGTCACTGTTTATGCTGGGATCTCCGCTCCGCAACCTGATATGCTTACATATG 804
QY 781 AAGAGATGACATGCTTTAGATGAAAGCTTACGATTTGTGAAAGAAAGAGCTTACT 840
DB 805 AAGAGATGACATGCTTTAGATGAAAGCTTACGATTTGTGAAAGAAAGAGCTTACT 864
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCCTGAGCTATGAGAGAGATTAGAAC 900
DB 865 ATATCTCCAACTTCAATTTTCTGGGCCAACTCCTGAGCTATGAGAGAGATTAGAAC 924
QY 901 CAGACT 906
DB 925 CAGACT 930

RESULT 6
AX180875 2732 bp DNA linear PAT 06-AUG-2001
LOCUS
DEFINITION
ACCESSION
VERSION
AX180875.1 GI:15132703
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 Plozman, G.D., Martinez, R., Whyte, D., Manning, G., Sudarshanam, S.,
Hill, R.V. and Flanagan, P.
TITLE
Mammalian protein phosphatases
JOURNAL
Patent: WO 0146394-A 2 28-JUN-2001;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
1..2732
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.6%; Score 902.8; DB 6; Length 2732;
Best Local Similarity 99.8%; Pred. No. 4, 8e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAAGTCAAAATTTGTTACTGAGAGGTTGCTCTGCTGAA 60
DB 538 ATGGCCCATGAGATGATGGAAGTCAAAATTTGTTACTGAGAGGTTGCTCTGCTGAA 597
QY 61 AGTGAACGAAAAAGTGTCTGCTTAATTTGATAGCCGCCCATTTGTGAAATTCATACATCC 120
DB 598 AGTGAACGAAAAAGTGTCTGCTTAATTTGATAGCCGCCCATTTGTGAAATTCATACATCC 657
QY 121 CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAGAGGTTGCAACAG 180
DB 658 CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAGAGGTTGCAACAG 717
QY 181 GACAAAGTGTAAATTAACAGAGCTATCCAGCATTCAGCGAAATAGATTGACATTGAT 240
DB 718 GACAAAGTGTAAATTAACAGAGCTATCCAGCATTCAGCGAAATAGATTGACATTGAT 777
QY 241 TGCAGTCAAGAGTGTATGTTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
DB 778 TGCAGTCAAGAGTGTATGTTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 837
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAGAGCTTCAACTCTGTTCACTG 360
DB 838 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAGAGCTTCAACTCTGTTCACTG 897
QY 361 CTTCAGAGTGGGTTTCTGAGATCTCTGTTGTTTCCCTGAGCTCTGTGAGAAATTC 420
DB 898 CTTCAGAGTGGGTTTCTGAGATCTCTGTTGTTTCCCTGAGCTCTGTGAGAAATTC 957
QY 421 ACTCTAGTCCCTACCTGATCTTCTGAGCTTGTCTGCTTACCTGTGCAACATTTGGCCCAACC 480
DB 958 ACTCTAGTCCCTACCTGATCTTCTGAGCTTGTCTGCTTACCTGTGCAACATTTGGCCCAACC 1017
QY 481 CGAATCTTCCCACTCTTATCTTGGCTGCGAGAGATGCTCTCAACAAGAGCTGATA 540
DB 1018 CGAATCTTCCCACTCTTATCTTGGCTGCGAGAGATGCTCTCAACAAGAGCTGATA 1077
QY 541 CAGCAGATGGGATGGTTATGTTTAAATGCGAGTATACCTGTCCAAAGCTGACTTT 600
DB 1078 CAGCAGATGGGATGGTTATGTTTAAATGCGAGTATACCTGTCCAAAGCTGACTTT 1137
QY 601 ATCCCGAGTCTCATTTCTGCGGTGCTGTGAAATGACAGCTTTTGTGAAATTTTG 660
DB 1138 ATCCCGAGTCTCATTTCTGCGGTGCTGTGAAATGACAGCTTTTGTGAAATTTTG 1197
QY 661 CCGTGTGTGACAAATCATGATGATTTCAATGAGAAAGCAAAAGCTTCAATGATGTT 720
DB 1198 CCGTGTGTGACAAATCATGATGATTTCAATGAGAAAGCAAAAGCTTCAATGATGTT 1257
QY 721 CTAGTCACTGTTTATGCTGGGATCTCCGCTCCGCAACCTGCTATGCTTACATATG 780
DB 1258 CTAGTCACTGTTTATGCTGGGATCTCCGCTCCGCAACCTGCTATGCTTACATATG 1317
QY 781 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAGAGCTTACT 840
DB 1318 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAGAGCTTACT 1377
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCCTGAGCTATGAGAGAGATTAGAAC 900
DB 1378 ATATCTCCAACTTCAATTTTCTGGGCCAACTCCTGAGCTATGAGAGAGATTAGAAC 1437
QY 901 CAGACT 906
DB 1438 CAGACT 1443

RESULT 7
AX278461 3059 bp DNA linear PAT 02-NOV-2001
LOCUS
DEFINITION
Sequence 1 from Patent WO0177340.

ACCESSION AX278461
VERSION AX278461.1 GI:16605915
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
Ducker, K.
Identification of a dual specificity phosphatase: dusp-10
Patent: WO 017340-A 1 18-Oct-2001;
MERCK PATENT GmbH (DE)
FEATURES
source Location/Qualifiers
1..3059
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
127..2124
/note="unamed protein product"
/codon_start=1
/protein_id="CAD10545.1"
/db_xref="GI:16605916"
/translation="MAHEMIGIOYIYERLVALLESCTEKLLIDISREPVYNTSHILE
AIVNCSKIMKRLLODKLITBELIQAHSKRVNDISQKVVVDOSQOVASLSDC
FLTVLGLKEKSPNSVHLIAGSFBSFQGLCEKSLVPTCISQCLPVANIGPT
ILPMLDKSVDFLEKASASNGCVLVHICLAGISRSATIAIAYIMKRMMSIDENRYVE
KRPETISPNFNFLGOLLDYEKKIKNOTGASGPKSLKLHLIEXNEPSPHVSBSQKSE
TPLSFPCADSAISEAAGORPVHPASVSPSPVOPSLDEPIYQALSGVHLSADRLD
SNLKRSEFLDKSVSASMAASIHGFSSEDELYKRSSTLLDSTINKLCPSPVOE
LSEQTPETSPPKEEASIPKRLQTPRPSDOSKRLHSYRISSTAGRSLLSPHRSQS
VEDNHTSFLEFLSTSOOHLTKSAGLKHWSHIDLAPOSTSLTSLSYFATESSHF
YSASAIVGSAVSAYSCSCLPTCGDDVIVSRROKRDADRSRHSRHSPEKQFK
RRSCOMERGESIMENSRREBELKRVSSQSSFSGSMETIEVS"

CDS

ORIGIN

Query Match 99.6%; Score 902.8; DB 6; Length 3059;
Best Local Similarity 99.8%; Pred. No. 4.9e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
1 ATGCCCATGATGATGATGGAATCAATTTGTAAGTGGAGGTTGGCTGCTGGAA 60
127 ATGCCCATGATGATGATGGAATCAATTTGTAAGTGGAGGTTGGCTGCTGGAA 186
61 AGTGAACGAAAAAGTCTGCTAATTTGTAAGTGGAGGTTGGCTGCTGGAA 120
187 AGTGAACGAAAAAGTCTGCTAATTTGTAAGTGGAGGTTGGCTGCTGGAA 246
121 CACATTTGGAGCCATTAATATCAATGCTGCTCAAGCTTATGAAGCGAGTTGCAAC 180
247 CACATTTGGAGCCATTAATATCAATGCTGCTCAAGCTTATGAAGCGAGTTGCAAC 306
181 GACAAAGTGTAAATTAAGAGCTCATCCAGATTCAGGAAACATAAGTTGACATTGAT 240
307 GACAAAGTGTAAATTAAGAGCTCATCCAGATTCAGGAAACATAAGTTGACATTGAT 366
241 TCGAGTCAGAGGTTGTAGTTACGATCAAGCTCCCAAGATTTGCTCTCTCTTCA 300
367 TCGAGTCAGAGGTTGTAGTTACGATCAAGCTCCCAAGATTTGCTCTCTCTTCA 426
301 GACTTTTTCACATGTAATCTTCTGGGTAACTGGAGAGAGCTTCAACTCTGTAACCTG 420
427 GACTTTTTCACATGTAATCTTCTGGGTAACTGGAGAGAGCTTCAACTCTGTAACCTG 486
361 CTTCAGAGTGGGTTGCTGAGATCTCTGTTTCCCTGGGCTCTGTAAGAGAAATCC 420
487 CTTCAGAGTGGGTTGCTGAGATCTCTGTTTCCCTGGGCTCTGTAAGAGAAATCC 546
421 ACTCTAGTCCCTACCTGCAATTTCTGAGCTTGTACCTGTTGCCAATATGGGCAACC 480
547 ACTCTAGTCCCTACCTGCAATTTCTGAGCTTGTACCTGTTGCCAATATGGGCAACC 606
```

481 CGAATTTCTCCCAATCTTTATCTTGGCTGCTCCAGCGAGATGCTTCAACAGAGCTGATA 540
607 CGAATTTCTCCCAATCTTTATCTTGGCTGCTCCAGCGAGATGCTTCAACAGAGCTGATA 540
541 CACAGAGATGAGATGTTGTTATGTTAAATGCGACTATACCTGCTCAAGCTGACTTT 600
667 CACAGAGATGAGATGTTGTTATGTTAAATGCGACTATACCTGCTCAAGCTGACTTT 726
601 ATCCCGAGTCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
727 ATCCCGAGTCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
661 CCGTGGTGGACAAATCAGTATGATTTGATTTGAGAGAGAGAGAGAGAGAGAGAGAG 786
787 CCGTGGTGGACAAATCAGTATGATTTGATTTGAGAGAGAGAGAGAGAGAGAGAGAG 846
721 CTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 846
847 CTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 906
781 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 906
907 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 966
841 ATATCTCCAACTTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
967 ATATCTCCAACTTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
901 CAGACT 906
1027 CAGACT 1032

RESULT 8
LOCUS BC042101
DEFINITION Homo sapiens dual specificity phosphatase 16, mRNA (cDNA clone
ACCESSION BC042101
VERSION BC042101.1 GI:27469788
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 3284)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schettler, T.E., Brownstein, M.J., Udell, T.B., Loquellano, N.A., Peters, G.J., Abramson, R.D., Malek, J.A., Bosak, S.A., McMan, P.J., McKernan, K.J., Wille, J., S.J., Bosak, S.A., Loquellano, N.A., Peters, G.J., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whiting, M., Kettelman, W., Madan, A., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, J.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Schermer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 3284)
Strausberg, R.
Direct Submission
JOURNAL MEDLINE
PUBMED 12477932
REFERENCE 22388257
AUTHORS Strausberg, R.
TITLE Submitted (23-DEC-2002) National Institutes of Health, Mammalian

REMARK
COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) md@paxll.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 88 Row: a Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction.

FEATURES
source

Location/Qualifiers
1. .3284
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:50665 IMAGE:4400399"
/issue_type="Duodenum, adenocarcinoma"
/clone_lib="NIH MGC_88"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .3284
/gene="DUSP16"
/note="Synonyms: KIAA1700, MKP-7, MKP7"
/db_xref="LocusID:80824"
/db_xref="MIM:607175"

CDS

332..2320
/codon_start=1
/product="DUSP16 protein"
/protein_id="AAH42101.1"
/db_xref="GI:27469789"
/translation="MAHMICTQIVTEPLVALLESTERYLLIDSPRYEYNTSHLE
AIIINCSKLMKRRILQODKVLITBLIHSKHKYIDCSQVYVYDSSQDVASLSDDC
PLTVLGLKLEKSPNSVHLAAGFAEFSRCPGLCEGKSLVPTCLISOPCLPVANIGPT
RIPLNVLGCGORDVNLKELMOQWIGVYVNASNTCPKDPILPEKSHLRVPVDSFCEK
ILPWLDSYDFLEKAKASNGCVLHCLAGISRSATIAIYIMKRMMSLDEAVRFYKE
KRPITSPNPFGLGLIDYEEKIKNOTGASGPKSKLHLKEMNEVPVAVSGGOKSE
TPLSPCCASATSEAGQRPVHPASVPSVQPSLEPVLVQALSGHLASDLRLDSNK
LKRSFELDIKSVYSASMAASHGFSSSDALRYKPTTLIDGTNLCPFSVQELISE
QTPETSPDKKEASIPKILQTPARSDQSRLHSVRTSSGTQORSLSLPSHSGVED
NYHTSEFLGLSTSQOHLTKSAGLGLKMSHDILAPQTSPTLSWSYFATESHFEVSA
SAIYGSASAYSASCSQLPTCGDQVYSVRROKPSRADRSRSHBESPFKEQFKRS
COMEFESIMSENRSEELGKVGSGSSPSFGSMELIIVS"
380..730
/note="RHOD; Region: Rhodanese Homology Domain"
/db_xref="CDD:smart00450"
803..1222
/note="DSPC; Region: Dual specificity phosphatase,
catalytic domain. Ser/Thr and Tyr protein phosphatases.
The enzyme's tertiary fold is highly similar to that of
tyrosine-specific phosphatases, except for a 'recognition'
region"
/db_xref="CDD:pfam00782"

ORIGIN

Query Match 99.6%; Score 902.8; DB 9; Length 3284;
Best Local Similarity 99.8%; Pred. No. 4,9e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 ATGGCCCATGAGATGATGGAAGTCAATTTGTTACTGAGAGTTGGTGGCTGTCTGAA 60
|||||

Db 332 ATGGCCCATGAGATGATGGAAGTCAATTTGTTACTGAGAGTTGGTGGCTGTCTGAA 391
QY 61 AGTGAACGGAAAAAGTGTCTGCTAATTGATAGCGGGCATTTGTGAAATACATATCC 120
Db 392 AGTGAACGGAAAAAGTGTCTGCTAATTGATAGCGGGCATTTGTGAAATACATATCC 451
QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAACTTATGAGAGGAAGTTGCAACAG 180
Db 452 CACATTTTGAAGCCATTATATCACTGCTCCAACTTATGAGAGGAAGTTGCAACAG 511
QY 181 GACAAAGTTTATATCAAGAGCTCATCAGCATTCAGGAAACATAAGTTGACATTGAT 240
Db 512 GACAAAGTTTATATCAAGAGCTCATCAGCATTCAGGAAACATAAGTTGACATTGAT 571
QY 241 TGCAGTCAGAGTTTATGATTAACATCAAGTCCCAAGATTTTCCCTCTCTTCA 300
Db 572 TGCAGTCAGAGTTTATGATTAACATCAAGTCCCAAGATTTTCCCTCTCTTCA 631
QY 301 GACTGTTTCTCACTGCTACTTCTGGGTTAACTGAGAGAGCTTCAACTCTTCACTG 360
Db 632 GACTGTTTCTCACTGCTACTTCTGGGTTAACTGAGAGAGCTTCAACTCTTCACTG 691
QY 361 CTTCAGAGTGGGTTTGGCTGAGTCTCTCGTGTGTTTCCCTGACCTCTGTGAGGAAATCC 420
Db 692 CTTCAGAGTGGGTTTGGCTGAGTCTCTCGTGTGTTTCCCTGACCTCTGTGAGGAAATCC 751
QY 421 ACTCTAGTCCCTACCTGCTGCTATTCAGGCTTGTACCTGTTGCCAATTTGGGCCAACC 480
Db 752 ACTCTAGTCCCTACCTGCTGCTATTCAGGCTTGTACCTGTTGCCAATTTGGGCCAACC 811
QY 481 CGAATCTTCCCAATCTTATCTGAGCTGCAAGCAGATGCTCTCAACAGAGAGTGAAT 540
Db 812 CGAATCTTCCCAATCTTATCTGAGCTGCAAGCAGATGCTCTCAACAGAGAGTGAAT 871
QY 541 CAGCAGAAATGGATGGTATATGTTAAATGCCAGCTATACCTGTCMAAGCTGACTTT 600
Db 872 CAGCAGAAATGGATGGTATATGTTAAATGCCAGCAATACCTGTCMAAGCTGACTTT 931
QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAAATGACAGCTTTGTGAGAAATTTTG 660
Db 932 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAAATGACAGCTTTGTGAGAAATTTTG 991
QY 661 CCGTGTGTGACAAATCAGTAGATTCATTGAGAAAGCAAAAGCTCCATGAGATGTT 720
Db 992 CCGTGTGTGACAAATCAGTAGATTCATTGAGAAAGCAAAAGCTCCATGAGATGTT 1051
QY 721 CTAGTCACTGTTTACCTGGATCTCCCGCTCGCCACCATGCTATCGCTACATCATG 780
Db 1052 CTAGTCACTGTTTACCTGGATCTCCCGCTCGCCACCATGCTATCGCTACATCATG 1111
QY 781 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTACT 840
Db 1112 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTACT 1171
QY 841 ATATCTCCAACTCAATTTTCTGGGCAACCTCTGACATGAGAGAAAGATTAGAAG 900
Db 1172 ATATCTCCAACTCAATTTTCTGGGCAACCTCTGACATGAGAGAAAGATTAGAAG 1231
QY 901 CAGACT 906
Db 1232 CAGACT 1237

RESULT 9
AX441210 3496 bp DNA linear PAT 28-JUN-2002
LOCUS AX441210
DEFINITION Sequence 1 from Patent WO0226997.
ACCESSION AX441210
VERSION AX441210.1 GI:21665766
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS
1 Lucie, R.M. and Wei, B.
TITLE
Dep-16 dual-specificity phosphatase
JOURNAL
Patent: WO 0226997-A 1 04-APR-2002;
Cephr, Inc. (US)

FEATURES
Source
Location/Qualifiers
1..3496
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match
Best Local Similarity 99.6%; Score 902.8; DB 6; Length 3496;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 1 ATGCCCCATGAGATGATTTGGAACCAATTTGTTACTGAGAGGTTGGCTCTGCTGGA 60
Db 562 ATGCCCCATGAGATGATTTGGAACCAATTTGTTACTGAGAGGTTGGCTCTGCTGGA 621
Oy 61 AGTGAACGGAAGAAAGTCTGCTTAATGATAGCCGCAATTTGTTGGAATACAAATACATCC 120
Db 622 AGTGAACGGAAGAAAGTCTGCTTAATGATAGCCGCAATTTGTTGGAATACAAATACATCC 681
Oy 121 CAATTTTGAAGCCATTAATATCAACTGCTCCAGCTTATGAAGCAAGGTTGCAACAG 180
Db 682 CAATTTTGAAGCCATTAATATCAACTGCTCCAGCTTATGAAGCAAGGTTGCAACAG 741
Oy 181 GACAAAGTGTAAATACAGAGCTCAAGCAATTCAGCAAGCAATTAAGTTGATGATGAT 240
Db 742 GACAAAGTGTAAATACAGAGCTCAAGCAATTCAGCAAGCAATTAAGTTGATGATGAT 801
Oy 241 TGCATCTGAAAGTTGATGATTAAGATCAAGCTCCCAAGATTTGCTCTCTCTTCA 300
Db 802 TGCATCTGAAAGTTGATGATTAAGATCAAGCTCCCAAGATTTGCTCTCTCTTCA 861
Oy 301 GACTGTTTCTCAGCTGATCTTGGGTAACCTGGAAGAGCTTCAACTGTTGACCTG 360
Db 862 GACTGTTTCTCAGCTGATCTTGGGTAACCTGGAAGAGCTTCAACTGTTGACCTG 921
Oy 361 CTTCAGAGTGGTGTGCTGATGTTCTCTGTTGTTTCCGAGCTCTGGAAGAAATTC 420
Db 922 CTTCAGAGTGGTGTGCTGATGTTCTCTGTTGTTTCCGAGCTCTGGAAGAAATTC 981
Oy 421 ACTCTAGTCCCTACCTGATCTTCTGAGCTTCTTACTGTTGCAACATTTGGCCAAAC 480
Db 982 ACTCTAGTCCCTACCTGATCTTCTGAGCTTCTTACTGTTGCAACATTTGGCCAAAC 1041
Oy 481 CGAATTCCTCCCAATCTTTATCTTGGCTGCAAGAGTCTTCAAGAGAGCTGATG 540
Db 1042 CGAATTCCTCCCAATCTTTATCTTGGCTGCAAGAGTCTTCAAGAGAGCTGATG 1101
Oy 541 CAGCAAGATGGATTTGATGTTAAATCCGCTATACCTGTTCCAAAGCTGACTT 600
Db 1102 CAGCAAGATGGATTTGATGTTAAATCCGCTATACCTGTTCCAAAGCTGACTT 1161
Oy 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGATGATGAGCTTGTGGAAGAAATTTTG 660
Db 1162 ATCCCGAGTCTCAATTTCTGCTGCTGCTGATGATGAGCTTGTGGAAGAAATTTTG 1221
Oy 661 CCGTGTGTGAAGCAATGATGATTTCAATGAGCAAGAAAGCTCCCAATGATGATG 720
Db 1222 CCGTGTGTGAAGCAATGATGATTTCAATGAGCAAGAAAGCTCCCAATGATGATG 1281
Oy 721 CTATGTCATGTTTAAAGTGTGCTCCGCTCCGCAACCATGCTATAGCTATAGTATG 780
Db 1282 CTATGTCATGTTTAAAGTGTGCTCCGCTCCGCAACCATGCTATAGCTATAGTATG 1341
Oy 781 AAGAGATGAGCATGCTTTAGATGAGCTTACAGATTTGGAAGAAAGAAAGCACTACT 840
Db 1342 AAGAGATGAGCATGCTTTAGATGAGCTTACAGATTTGGAAGAAAGAAAGCACTACT 1401

```

Oy 841 ATATCTCAAACTGCAATTTTCTGGCCCACTCTGATGATGAGAGAAATTAAAC 900
Db 1402 ATATCTCAAACTGCAATTTTCTGGCCCACTCTGATGATGAGAGAAATTAAAC 1461
Oy 901 CAGACT 906
Db 1462 CAGACT 1467

RESULT 10

AB052156 3521 bp mRNA linear PRI 18-OCT-2001
DEFINITION
Homo sapiens MKP-7 mRNA for MAPK phosphatase-7, complete cde.
AB052156.1 GI:13548676
KEYWORDS

SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

AUTHORS
1 Masuda, K., Shima, H., Watanabe, M. and Kikuchi, K.
TITLE
MKP-7, a novel mitogen-activated protein kinase phosphatase,
functions as a shuttle protein
JOURNAL
U. Biol. Chem. 276 (42), 39002-39011 (2001)
MEDLINE
21486429
PUBMED
11489891

REFERENCE
AUTHORS
1 Masuda, K., Shima, H. and Kikuchi, K.
TITLE
Direct Submission
Submitted (05-DEC-2000) Kouhei Masuda, Institute for Genetic
Medicine, Division of Biochemical Oncology and Immunology, Kita-ku
Kita15-Jou nishi7-Eyome, Sapporo, Hokkaido 060-0815, Japan
(E-mail: kou@imm.hokudai.ac.jp, Tel:81-11-706-5536,
Fax:81-11-707-6839)

FEATURES

source

Location/Qualifiers

1..3521
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/cell_line="Jurkat"
1..3521
/gene="MKP-7"
564..-2561
/gene="MKP-7"
/codon_start=1
/product="MAPK phosphatase-7"
/protein_id="BAB40814.1"
/translation="MAHEMTGTOIVTERLVALLSEETKYLIDSRPEVYNTSHILE
AATNCSKLMRRLQODKVLITELIQSAHAKVDIOSKQVWVDSQDVASLSDC
FLTVLGLKESFNSVHLIAGPFRSFCPLGCEKSTLVPCISQPCLPVANTGPT
ILPMLVYGCORVDIANKELMOQNGIGVLAASNTCPKPIPESHFLRVVDSRCEK
KRPTISPRNFTGLDYERKINQOTGASGPKSLKLHLIEKRPVPAVSQGGKSE
TLPSPCADSTSEAGQRPVHPASVPSVOPSLIEDSPVOLSGLHLSADLED
SNKRSFSDIDISVSYSASMAASHGPFSSSDALVYKPSPTLLGTNKLQFSPVDE
LSEOTPEPTSPDKESASIPKKLTOTARPDSOSRSLSVYSSGTARQSLPLRSGS
VSAIATGYSFLFGLSTSOHLTKSAGLGLQMSDILAQSTSPSLTSSVYATSSSF
RSCQMFSGSINSRSEBELGKVSGSSFSGSEIIEVS"

gene

CDS

ORIGIN

Query Match
Best Local Similarity 99.6%; Score 902.8; DB 9; Length 3521;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 1 ATGCCCCATGAGATGATTTGGAACCAATTTGTTACTGAGAGGTTGGCTCTGCTGGA 60
Db 564 ATGCCCCATGAGATGATTTGGAACCAATTTGTTACTGAGAGGTTGGCTCTGCTGGA 623
Oy 61 AGTGAACGGAAGAAAGTCTGCTTAATGATAGCCGCAATTTGTTGGAATACAAATACATCC 120

```


Db 624 AGTGAACGAAAGAGCTGCTATGATAGCGGCACTTTGTGGAATACATACATCC 663
 Qy 121 CACATTTTGGAGCCATTAAATATCACTGCTCCAGCTTATGAAGGAGGTTGCAACAG 180
 Db 684 CACATTTTGGAGCCATTAAATATCACTGCTCCAGCTTATGAAGGAGGTTGCAACAG 743
 Qy 181 GACAAAGTGTATTAATTAAGAGCTCATCAGCATTCAGCGAAACATAAGTGAATTGAT 240
 Db 744 GACAAAGTGTATTAATTAAGAGCTCATCAGCATTCAGCGAAACATAAGTGAATTGAT 803
 Qy 241 TGGAGTCAGAGGTTGATGTTATGATCAAGAGCTCCAGAGATGTCCTCTCTCTTCA 300
 Db 804 TGGAGTCAGAGGTTGATGTTATGATCAAGAGCTCCAGAGATGTCCTCTCTCTTCA 863
 Qy 301 GACTGTTTCTCACTGATCTTCTGGGTAACCTGGAGAGGCTTCAACTCTGTTCACTG 360
 Db 864 GACTGTTTCTCACTGATCTTCTGGGTAACCTGGAGAGGCTTCAACTCTGTTCACTG 923
 Qy 361 CTGCAAGTGGGTTTGTGATGTTCTGCTGTTTCCCTGGCTCTGTTGAAGGAAATCC 420
 Db 924 CTGCAAGTGGGTTTGTGATGTTCTGCTGTTTCCCTGGCTCTGTTGAAGGAAATCC 983
 Qy 421 ACTCTAGTCCCTACCTGCACTTCTCAAGCTTGTCTTACCTGTCACAACTTGGGCCAAC 480
 Db 984 ACTCTAGTCCCTACCTGCACTTCTCAAGCTTGTCTTACCTGTCACAACTTGGGCCAAC 1043
 Qy 481 CGAATTTCTCCAACTTTATCTTGTGCTGCGCAAGAGATGTCCTCAAGAGAGCTGATA 540
 Db 1044 CGAATTTCTCCAACTTTATCTTGTGCTGCGCAAGAGATGTCCTCAAGAGAGCTGATA 1103
 Qy 541 CAGCAGAGTGGATGTTGTTATGTTAAATGCGAGATGTCCTGTCAGAGCTGATG 600
 Db 1104 CAGCAGAGTGGATGTTGTTATGTTAAATGCGAGATGTCCTGTCAGAGCTGATG 1163
 Qy 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTAATGACAGCTTGTGTGAGAAATTTTG 660
 Db 1164 ATCCCGAGTCTCATTTCTGCGTGTGCTGTAATGACAGCTTGTGTGAGAAATTTTG 1223
 Qy 661 CCGTGTGGAGCAATGATGATTTATTTAGAGAAAGCAAGCTTCAATGATGTTG 720
 Db 1224 CCGTGTGGAGCAATGATGATTTATTTAGAGAAAGCAAGCTTCAATGATGTTG 1283
 Qy 721 CTAGTGAAGTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 780
 Db 1284 CTAGTGAAGTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 1343
 Qy 781 AAGAGATGAGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATG 840
 Db 1344 AAGAGATGAGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATG 1403
 Qy 841 ATATCTCAAGTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 900
 Db 1404 ATATCTCAAGTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 1463
 Qy 901 CAGACT 906
 Db 1464 CAGACT 1469

RESULT 11
 AX260340 3544 bp DNA linear PART 26-OCT-2001
 LOCUS AX260340
 DEFINITION Sequence 1 from Patent WO0173059.
 ACCESSION AX260340
 VERSION AX260340.1 GI:16509303
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Meyers, R.A.

TITLE 38692 and 21117: dual specificity phosphatase molecules and uses
 JOURNAL Patent: WO 0173059-A 1 04-OCT-2001.
 FEATURES
 source Millennium Pharmaceuticals, Inc. (US)
 Location/Qualifiers
 1..3544
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 589..2586
 /note="unassigned protein product"
 /protein_id="CAD10218.1"
 /codon_start=1
 /db_xref="GI:16509304"
 /db_xref="RefSeq:NM_001021818"
 /translation="MAHEMIGTIVTERVALLESTETKYLIDSPPEVYNTSHLE
 AININCSKLMKRLQODKYLITELIQHSAKHKRDVDSQKVVYDSDSDVALSDDC
 FLTVLLGKLEKSFNSVHLAAGFAEBSRPGICEKSLVPCISQPCLPVANIAPT
 RILPMLYLCORDVNLKELMOQNGIGVYVNASNTCPKPPFIPESHFLRVVNDSPCEK
 ILPMLDYSVDPIEKAKASNGCVLHGLAGISRSATIAIYIMRMMSIDEAYRFXE
 KPTTSPNPNFAGQILDYERKIKNGTGAAGRPKSLILHLEKNEVPVAVSEGGQSE
 TPSPPCADSAISEAAGRPVHPASVPVSVQPSLLEDSPLVQALSGHLNLSADRLSD
 SNLKRKSFSLDIKSVYSASMAASLHGFSSSEDLAEYKRSITLDTGNTLQPSVQVE
 LSEQTEPSPDKEASIPKLTQARPSDSQSKRLHSEVTSSEGTQARSLISPLHRSQS
 VEDNHTSFLPGISTSOQHLTKTSAGLGMHSDILAPQSTSLTSWYFAESSHF
 YASATYGGSAYSAYSASOLPCCGPOVYSVVRKRPDPSRADRSRMHBSPREKQFK
 RNSCMERFESITSENRSRELEKVGSGSFSQSMETIEVS"

ORIGIN
 Query Match 99.6%; Score 902.8; DB 6; Length 3544;
 Best Local Similarity 99.8%; Pred. No. 5e-267;
 Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGCCCCATGAGATGATGTTGAACTCAATTTGTTACTGAGAGGTTGGCTGCTGGAA 60
 Db 589 ATGCCCCATGAGATGATGTTGAACTCAATTTGTTACTGAGAGGTTGGCTGCTGGAA 648
 Qy 61 AGTGAACGAAAGAGCTGCTATGATAGCGGCACTTTGTGGAATACATACATCC 120
 Db 649 AGTGAACGAAAGAGCTGCTATGATAGCGGCACTTTGTGGAATACATACATCC 708
 Qy 121 CACATTTTGGAGCCATTAAATATCACTGCTCCAGCTTATGAAGGAGGTTGCAACAG 180
 Db 709 CACATTTTGGAGCCATTAAATATCACTGCTCCAGCTTATGAAGGAGGTTGCAACAG 768
 Qy 181 GACAAAGTGTATTAATTAAGAGCTCATCAGCATTCAGCGAAACATAAGTGAATTGAT 240
 Db 769 GACAAAGTGTATTAATTAAGAGCTCATCAGCATTCAGCGAAACATAAGTGAATTGAT 828
 Qy 241 TGGAGTCAGAGGTTGATGTTATGATCAAGAGCTCCAGAGATGTCCTCTCTCTTCA 300
 Db 829 TGGAGTCAGAGGTTGATGTTATGATCAAGAGCTCCAGAGATGTCCTCTCTCTTCA 888
 Qy 301 GACTGTTTCTCACTGATCTTCTGGGTAACCTGGAGAGGCTTCAACTCTGTTCACTG 360
 Db 889 GACTGTTTCTCACTGATCTTCTGGGTAACCTGGAGAGGCTTCAACTCTGTTCACTG 948
 Qy 361 CTGCAAGTGGGTTTGTGATGTTCTGCTGTTTCCCTGGCTCTGTTGAAGGAAATCC 420
 Db 949 CTGCAAGTGGGTTTGTGATGTTCTGCTGTTTCCCTGGCTCTGTTGAAGGAAATCC 1008
 Qy 421 ACTCTAGTCCCTACCTGCACTTCTGCGTGTGCTGTTACCTGTTGCAACATTTGGGCCAAC 480
 Db 1009 ACTCTAGTCCCTACCTGCACTTCTGCGTGTGCTGTTACCTGTTGCAACATTTGGGCCAAC 1068
 Qy 481 CGAATTTCTCCAACTTTATCTTGTGCTGCGCAAGAGATGTCCTCAAGAGAGCTGATA 540
 Db 1069 CGAATTTCTCCAACTTTATCTTGTGCTGCGCAAGAGATGTCCTCAAGAGAGCTGATA 1128
 Qy 541 CAGCAGAGTGGATGTTGTTATGTTAAATGCGAGATGTCCTGTCAGAGCTGATG 600
 Db 1129 CAGCAGAGTGGATGTTGTTATGTTAAATGCGAGATGTCCTGTCAGAGCTGATG 1188

Query	Match	Best Local Similarity	99.8%;	Score 902.8;	DB 9;	Length 3566;
Matches	904;	Conservative	0;	Mismatches	2;	Indels 0; Gaps 0;
1	ATGSCCCATGAGATTGGAACCTCAATTGTCTGAGAGCTTGCGCTCTGCTGGAA	60				
633	ATGGCCCATGAGATTGGAACCTCAATTGTCTGAGAGCTTGCGCTCTGCTGGAA	692				
61	AGTGAACGGAAAAAGTGTGCTGAATTGATACCGCGCAATTGTGGAATACATACATCC	120				
693	AGTGAACGGAAAAAGTGTGCTGAATTGATACCGCGCAATTGTGGAATACATACATCC	752				

Db	753	CACATTTTGAAAGCCATTAATATCAACAGCTCCAAAGCTTATGAAGCAGAAAGTTTGCAACAG	812
Oy	181	GACAAAGTTTAATTACAGAGCTATCCAGATTCAGCGAAACATAAGTTTGACATTGAT	240
Db	813	GACAAAGTTTAATTACAGAGCTATCCAGATTCAGCGAAACATAAGTTTGACATTGAT	872
Oy	241	TGCAGTCGAAAGTTTGATAGTTTACGATCAAAAGCTCCCAAGATGTTGACCTCTCTCTTCA	300
Db	873	TGCATTCGAAAGTTTGATAGTTTACGATCAAAAGCTCCCAAGATGTTGACCTCTCTCTTCA	932
Oy	301	GACTGTTTTCTCACTGACTCTGGGTAACATGGAAGAGCTTCAACTCTTGCACTG	360
Db	933	GACTGTTTTCTCACTGACTCTGGGTAACATGGAAGAGCTTCAACTCTTGCACTG	420
Oy	361	CTTGAGGTGGGTTTGCTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGGAAATTC	480
Db	993	CTTGAGGTGGGTTTGCTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGGAAATTC	1052
Oy	421	ACTGTAGTCCCTACCTGCATTTCTCAGAGCTTGCTTAACCTGTTGCCAACATTTGGCCAAAC	480
Db	1053	ACTGTAGTCCCTACCTGCATTTCTCAGAGCTTGCTTAACCTGTTGCCAACATTTGGCCAAAC	1112
Oy	481	CGAATCTTCCCATTTTATCTTGAGCTGCAGCGAGATGTCTCAACAGAGACTATA	540
Db	1113	CGAATCTTCCCATTTTATCTTGAGCTGCAGCGAGATGTCTCAACAGAGACTATG	1172

D _b	1233	ATCCCGGAGCTCAATTTCCTGGCGTGCCTGTGAATGACAGCTTTTGTAAGAATAATTTTG	1292
Q _y	661	CCGTGTTGGACAATCAGTAGAATTTCAATGAGAAGCAAAGCCCTCCATGATGTGTT	720
D _b	1233	CCGGTGTGGACAATCAGTAGAATTTCAATGAGAAGCAAAGCCCTCCATGATGTGTT	1352
Q _y	721	CTAGTGACATGTTTAACTGGGAGATCTCCGCTCGGCCACCATGCTATGCGCTACTCATATG	780
D _b	1353	CTATGTGACATGTTTAACTGGGAGATCTCCGCTCGGCCACCATGCTATGCGCTACTCATATG	1412
Q _y	781	AAGAGATGACATGCTTTAGATGAAGGTTACATTTTGGAAGAAAAAAGAACCTTACT	840
D _b	1413	AAGAGATGACATGCTTTAGATGAAGGTTACATTTTGGAAGAAAAAAGAACCTTACT	1472
Q _y	841	ATATCTCCAAACTCAATTTTCTGGGCCCACTCTCTGACCTATGAGAAGAAAGATTAAAGAC	900
D _b	1473	ATATCTCCAAACTCAATTTTCTGGGCCCACTCTCTGACCTATGAGAAGAAAGATTAAAGAC	1536

Query 901 CAGACT 906
Db 1533 CAGACT 1538

RESULT 13
LOCUS AX374994 3766 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 17 from Patent WO0210363.
ACCESSION AX374994
VERSION AX374994.1 GI:19169826
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Tang, Y.T., Elliott, V.S., Ramkumar, J., Yao, M.G., Burford, N., Wang, Y.E., Stewart, E.A., Gandhi, A.R., Patterson, C., Lee, E.A., Hafalla, A.J., Lu, D.A., Tribouley, C.M., Griffin, J.A., Baughn, M.R., Yue, H., Warren, B.A., Nguyen, D.B. and Walla, N.K.
TITLE Protein phosphatases
JOURNAL Patent: WO 0210363-A 17 07-FEB-2002;
Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1..3766
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7480570CB1"

ORIGIN
Query Match 99.6%; Score 902.8; DB 6; Length 3766;
Best Local Similarity 99.8%; Pred. No. 5e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGGCCCATGAGTATGAGTGAACCAATTGTTACTGAGAGTGGGCTGCTGGAGAA 60
538 ATGGCCCATGAGTATGAGTGAACCAATTGTTACTGAGAGTGGGCTGCTGGAGAA 597
61 AGTGAACGGAAGAGTGTCTGCTAATTTGATAGCGGCGCATTTGGATACAAATAC 120
598 AGTGAACGGAAGAGTGTCTGCTAATTTGATAGCGGCGCATTTGGATACAAATAC 657
121 CACATTTTGAAGCCATTAAATATCACTGCTCAAGTTATGAAGGAGTTGCAACAG 180
658 CACATTTTGAAGCCATTAAATATCACTGCTCAAGTTATGAAGGAGTTGCAACAG 717
181 GACAAAGTGTATTTTACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATT 240
718 GACAAAGTGTATTTTACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATT 777
241 TGCAGTCAGAGGTTGATTTAGATCAAGATCCCAAGATGTGCTCTCTCTTCA 300
778 TGCAGTCAGAGGTTGATTTAGATCAAGATCCCAAGATGTGCTCTCTCTTCA 837
301 GACTGTTTCTCACTGTAATCTTGGGTAACCTGAGAGAGCTTCAACTCTGTTCACT 360
838 GACTGTTTCTCACTGTAATCTTGGGTAACCTGAGAGAGCTTCAACTCTGTTCACT 897
361 CTTCGAGGTGGGTTGTGAGTCTCTCGTGTGTTCCCTGGGCTCTGTGAGAGAAATCC 420
898 CTTCGAGGTGGGTTGTGAGTCTCTCGTGTGTTCCCTGGGCTCTGTGAGAGAAATCC 957
421 ACTCTAGTCCCTACCTGATCTTCTGAGCTTGTGCTTACCTGTTCAACATTTGGGCA 480
958 ACTCTAGTCCCTACCTGATCTTCTGAGCTTGTGCTTACCTGTTCAACATTTGGGCA 1017
481 GGAATTTTCCCAATCTTTATCTTGGGTCGACGAGAGATGCTTCAACAGAGCTGATA 540
1018 GGAATTTTCCCAATCTTTATCTTGGGTCGACGAGAGATGCTTCAACAGAGCTGATA 1077
541 CACGAGATGGAGTTGGTTATGTATTAATGCGAGATACCTGTCCAAAGCTGACTTT 600

Db 1078 CAGCAGAAATGGAGTTGGTTATGTATTAATGCCAGCAATACCTGTCCAAAGCTGACTTT 1137
Query 601 ATCCCGAGTCTCATTTCTGCGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 660
Db 1138 ATCCCGAGTCTCATTTCTGCGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 1197
Query 661 CCGTGGTTGAGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCCAATGATGTGT 720
Db 1198 CCGTGGTTGAGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCCAATGATGTGT 1257
Query 721 CTATGCACTGTTTAACTGAGATCTCCGCTCCGCAACATGCTATGCTTACATCATG 780
Db 1258 CTATGCACTGTTTAACTGAGATCTCCGCTCCGCAACATGCTATGCTTACATCATG 1317
Query 781 AAGAGATGACATGCTTTTATGATGAGCTTACAGATTTGTGAGAAAGAAAGACTACT 840
Db 1318 AAGAGATGACATGCTTTTATGATGAGCTTACAGATTTGTGAGAAAGAAAGACTACT 1377
Query 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGATGAGAAAGATTAGAAC 900
Db 1378 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGATGAGAAAGATTAGAAC 1437
Query 901 CAGACT 906
Db 1438 CAGACT 1443

RESULT 14
BD171157 4790 bp DNA linear PAT 17-JAN-2003
LOCUS BD171157
DEFINITION Novel gene and protein encoded thereby.
ACCESSION BD171157
VERSION BD171157.1 GI:27876969
KEYWORDS WO 02052005-A/13.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Ohara, O., Nagase, T. and Nakajima, D.
TITLE Novel gene and protein encoded thereby
JOURNAL Patent: WO 02052005-A 13 04-JUL-2002;
KAZUSA, DNA RESEARCH INSTITUTE, OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
COMMENT OS Homo sapiens (human)
FN WO 02052005-A/13
PD 04-JUL-2002
PF 20-DEC-2001 WO 2001JP011217
PR 22-DEC-2000 JP 00P 389742
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/12,C07K14/47
CC Novel gene and protein encoded thereby
FH Key Location/Qualifiers
FT CDS (184)..(2178).

FEATURES
source Location/Qualifiers
1..4790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.6%; Score 902.8; DB 6; Length 4790;
Best Local Similarity 99.8%; Pred. No. 5.2e-267;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGGCCCATGAGTATGAGTGAACCAATTGTTACTGAGAGTGGGCTGCTGGAGAA 60
Db 184 ATGGCCCATGAGTATGAGTGAACCAATTGTTACTGAGAGTGGGCTGCTGGAGAA 243
Query 61 AGTGAACGGAAGAGTGTCTGCTAATTTGATAGCGGCGCATTTGGATACAAATAC 120
Db 244 AGTGAACGGAAGAGTGTCTGCTAATTTGATAGCGGCGCATTTGGATACAAATAC 303

QY	121	CACATTTTGGAGCCATTAATATCACTGCTCCAAAGCTTAATGAAACGAAAGTTGCAACG	180
Db	304	CACATTTTGGAGCCATTAATATCACTGCTCCAAAGCTTAATGAAACGAAAGTTGCAACG	363
QY	181	GACAAAGTGTAAATTCAGAGCTATCCAGCATTTAGCGAAACAATAAGTTGACATTGAT	240
Db	364	GACAAAGTGTAAATTCAGAGCTATCCAGCATTTAGCGAAACAATAAGTTGACATTGAT	423
QY	241	TGACGTCAGAAAGTGTAAATTTAGCATCAAAAGCTCCAAAGTGTGCTCTCTCTTCA	300
Db	424	TGACGTCAGAAAGTGTAAATTTAGCATCAAAAGCTCCAAAGTGTGCTCTCTCTTCA	483
QY	301	GACTGTTTTCTCACTGTACTCTGGGTAAATGGAAGAAAGCTTCAACTGTGTCACTG	360
Db	484	GACTGTTTTCTCACTGTACTCTGGGTAAATGGAAGAAAGCTTCAACTGTGTCACTG	543
QY	361	CTTGGAGGTGGGTTTGCCTGAGTTCTCTGTTGTTTTCCCTGGGCTCTGTAAAGAAATTC	420
Db	544	CTTGGAGGTGGGTTTGCCTGAGTTCTCTGTTGTTTTCCCTGGGCTCTGTAAAGAAATTC	603
QY	421	ACTCTAGTCCCTACCTGCAATTTCTCAAGCCTTGCTTACTGTTGCCAACTTGGGCCAAC	480
Db	604	ACTCTAGTCCCTACCTGCAATTTCTCAAGCCTTGCTTACTGTTGCCAACTTGGGCCAAC	663
QY	481	CGAATTTCTCCCAATCTTTATCTTGGCTGCCAGAGATGTCTCAACAAAGAGCTGATA	540
Db	664	CGAATTTCTCCCAATCTTTATCTTGGCTGCCAGAGATGTCTCAACAAAGAGCTGATA	723
QY	541	CAGCAGAAATGGATTTGGTTATGTGTAAATGCGACATTAACCTGTCCAAAGCCTGACTT	600
Db	724	CAGCAGAAATGGATTTGGTTATGTGTAAATGCGACAAATTAACCTGTCCAAAGCCTGACTT	783
QY	601	ATCCCGAGATCTCAATTTCTCGCGTGTGCTGTGATGATGACAGCTTTTGTGAGAAATTTTG	660
Db	784	ATCCCGAGATCTCAATTTCTCGCGTGTGCTGTGATGATGACAGCTTTTGTGAGAAATTTTG	843
QY	661	CCGTGTTGGACAAATCTAGATGATTTCAATGAGAAACAAAGCCTTCCAAATGAGATGTT	720
Db	844	CCGTGTTGGACAAATCTAGATGATTTCAATGAGAAACAAAGCCTTCCAAATGAGATGTT	903
QY	721	CTAGTGCACTGTTTAAAGCTGGGATCTCCCGCTCCGCCAACATCGCTATGCGCTACATCAG	780
Db	904	CTAGTGCACTGTTTAAAGCTGGGATCTCCCGCTCCGCCAACATCGCTATGCGCTACATCAG	963
QY	781	AAGAGATGACATGTCTTAAATGAGAGCTTACAGATTTGTGAAAGAAAAAGACTTACT	840
Db	964	AAGAGATGACATGTCTTAAATGAGAGCTTACAGATTTGTGAAAGAAAAAGACTTACT	1023
QY	841	ATAATCTCAAACTTCAATTTTCTGGGCCAACTCTCTGACATTAAGAGAAAGATTAAGAAC	900
Db	1024	ATAATCTCAAACTTCAATTTTCTGGGCCAACTCTCTGACATTAAGAGAAAGATTAAGAAC	1083
QY	901	CAGACT 906	
Db	1084	CAGACT 1089	

RESULT 15	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
BD183422	4790 bp	DNA	linear	PAT 17-JUN-2003			
BD183422		Novel genes and proteins encoded by the genes.					
BD183422.1	GI:31875622						
JP 2002345492-A/135							
Homo sapiens (human)							
Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.							
1 (bases 1 to 4790)							
Ohata,O., Nagase,T. and Nakajima, D.							
Novel genes and proteins encoded by the genes							
Patent: JP 2002345492-A 135 03-DEC-2002;							

COMMENT	FEATURES
OS Homo sapiens (human)	FT
PN JP 2002345492-A/135	CDS
PD 03-DEC-2002	Location/Qualifiers
PF 26-FEB-2002 JP 2002049009	1..4790
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA	Location/Qualifiers
PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00	1..4790
PC A61P25/14,	Location/Qualifiers
PC A61P25/18, A61P35/00, C12N15/00, A61K37/02	1..4790
CC Novel genes and proteins encoded by the genes FH	Location/Qualifiers
Location/Qualifiers	184) . (2178) .
source	1..4790
ORIGIN	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"

Query Match	99.6%	Score 902.8	DB 6	Length 4790
Best Local Similarity	99.8%	Pred. No. 5,2e-267		
Matches 904	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY	1	ATGCCCATGAGATGATGGAACTCAATTTGTAACTGAGAGGTTGGTCTCTCTGGAA	60	
Db	184	ATGGCCCATGAGATGATTTGGAACTCAATTTGTAACTGAGAGGTTGGTCTCTCTGGAA	60	
QY	61	AGTGAACGGAAAAAGTCTGCTAAATTGATATGACCGGCCCTTTGTGGATACAAATACCC	243	
Db	244	AGTGAACGGAAAAAGTCTGCTAAATTGATATGACCGGCCCTTTGTGGATACAAATACCC	303	
QY	121	CACATTTTGGAAAGCCATTAAATCAACTGCTCAAGCTTAAAGAACGAGTTGGCAACAG	180	
Db	304	CACATTTTGGAAAGCCATTAAATCAACTGCTCAAGCTTAAAGAACGAGTTGGCAACAG	363	
QY	181	GACAAAGTGTATATTAACAAGCTCAACGACATTCAGACATTCAGCAACATTAAGTTGACATTTGAT	240	
Db	364	GACAAAGTGTATATTAACAAGCTCAACGACATTCAGACATTCAGCAACATTAAGTTGACATTTGAT	423	
QY	241	TGCAGTCAGAAAGGTTGTATGTTTACATTAAGCTCCCAAGATGTGCTCTCTCTTCA	300	
Db	424	TGCAGTCAGAAAGGTTGTATGTTTACATTAAGCTCCCAAGATGTGCTCTCTCTTCA	483	
QY	301	GACTGTTTTCTACAGTATCTTCTGGGTAACTGAGAGAGCTCAACTCTGTATACCGT	360	
Db	484	GACTGTTTTCTACAGTATCTTCTGGGTAACTGAGAGAGCTCAACTCTGTATACCGT	543	
QY	361	CTTGAGAGTGGGTTTGCAGAGTTCTCTCGTGTGTTCCCGGCTCTGTGAGAAAAATCC	420	
Db	544	CTTGAGAGTGGGTTTGCAGAGTTCTCTCGTGTGTTCCCGGCTCTGTGAGAAAAATCC	603	
QY	421	ACTCAGAGCCCTACCTGATCTTCTACGCTTGTAACTCTGTCGCCAATTTGGGCAACC	480	
Db	604	ACTCAGAGCCCTACCTGATCTTCTACGCTTGTAACTCTGTCGCCAATTTGGGCAACC	663	
QY	481	CGAATTTCTTCCCAATCTTATCTTGGCTGTCAGCAGATGTCTCAACAGAGCTGATA	540	
Db	664	CGAATTTCTTCCCAATCTTATCTTGGCTGTCAGCAGATGTCTCAACAGAGCTGATA	723	
QY	541	CAGCAGATAGGATTTGATGTTAAATGCAAGCTATACCTGTCCAAAGCCCTGACTTT	600	
Db	724	CAGCAGATAGGATTTGATGTTAAATGCAAGCTATACCTGTCCAAAGCCCTGACTTT	783	
QY	601	ATGCCCGAATCTCATTTTCTGTGCTGTGCTGTGAATACAGAGCTTTGTGAGAAATTTTG	660	
Db	784	ATGCCCGAATCTCATTTTCTGTGCTGTGCTGTGAATACAGAGCTTTGTGAGAAATTTTG	843	
QY	661	CCGTGTTTGAACAATCGATGATTTCAATGAGAAAGCAAAAGCCCTCAATGATGATGTT	720	
Db	844	CCGTGTTTGAACAATCGATGATTTCAATGAGAAAGCAAAAGCCCTCAATGATGATGTT	903	
QY	721	CTAAGTCAGCTGTTTACGTGGGATCTCCCGCTCGCAGACATGGCTATGGCTTACATCATG	780	

```

Db      904 CTAGTGCACCTGTTAGCTGGGATCTCCCGCTCCGCCACATCGCTATCGCCTACATCATG 963
QY      781 AAGAGATGACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 840
Db      964 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 1023
QY      841 ATATCTCCAAACTTCAATTTCTGGGCCAACTCCTGGACTATGAGAAAGATTAAAGAAC 900
Db      1024 ATATCTCCAAACTTCAATTTCTGGGCCAACTCCTGGACTATGAGAAAGATTAAAGAAC 1083
QY      901 CAGACT 906
Db      1084 CAGACT 1089

```

Search completed: June 21, 2004, 23:40:22
 Job time : 5706.52 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 17:00:31 ; Search time 3586.84 Seconds
(without alignments)
7542.898 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 906
Sequence: 1 atggcccaatgagatgatg.....agaagataagaaccagact 906

Scoring table:
IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_esthum:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vtc:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vtl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	799.8	88.3	898	13	BQ721265
2	665.2	73.4	836	12	B1821804
3	633	69.9	682	12	B1816954
4	565.8	62.5	715	14	CF727177

5	559.6	61.8	639	14	CF742387	CF742387	UI-M-HB0-
6	558.2	61.6	566	14	CF532917	CF532917	UI-M-GH0-
7	551.4	60.9	769	13	BUT04078	BUT04078	UI-M-FOO-
8	544.8	60.1	650	10	BF472046	BF472046	UI-M-BH3-
9	541.4	59.8	602	10	AM847425	AM847425	RC1-CT020
10	535.6	59.1	595	10	AM847426	AM847426	RC1-CT020
11	489.6	54.0	862	13	BQ770036	BQ770036	UI-M-F10-
12	482.8	53.3	3335	11	AK035552	AK035552	Mus muscu
13	482.8	47.8	792	13	BU946566	BU946566	703769910
14	430.8	47.5	946	12	BI557528	BI557528	603281865
15	428.8	47.3	512	10	BF815601	BF815601	MR2-C1012
16	422.8	46.7	516	29	AY413655	AY413655	Homo sapi
17	415.8	45.9	1060	12	BM546940	BM546940	AGENCOURT
18	404.6	44.7	619	29	CG583869	CG583869	OST226476
19	399.6	44.1	463	10	CG058779	CG058779	naflf0604
20	390.8	43.1	903	12	BG482429	BG482429	602526595
21	386.4	42.6	463	10	BF809325	BF809325	QV0-C1020
22	373.2	41.2	516	29	AY413657	AY413657	Mus muscu
23	364.2	40.2	403	10	BE004625	BE004625	CM2-BN010
24	364.2	40.2	2207	11	BC038231	BC038231	Homo sapi
25	359.4	39.7	484	13	BQ318152	BQ318152	RC21-CT04
26	359.4	39.7	494	13	BQ318155	BQ318155	RC1-CT046
27	356.4	39.3	421	10	BE346884	BE346884	UI-M-BH3-
28	352.6	38.9	1148	12	BM906608	BM906608	AGENCOURT
29	352.2	38.9	924	14	CA474739	CA474739	AGENCOURT
30	349.6	38.6	404	10	BB816652	BB816652	BB816652
31	341	37.6	486	10	AM855006	AM855006	PM3-CT026
32	337.2	37.2	442	14	CB606149	CB606149	AMGNNUC:N
33	333.8	36.8	412	9	AA879894	AA879894	vw01d07.x
34	327.6	36.2	366	10	BG002947	BG002947	RC6-GN007
35	323.2	35.7	402	29	CG499793	CG499793	OST42180
36	322	35.5	685	14	CB165612	CB165612	BE0603161
37	318	35.1	579	10	BF582366	BF582366	602094304
38	293.6	32.6	324	13	BQ367284	BQ367284	RC4-GN022
39	293.6	32.4	635	10	BE377720	BE377720	601229866
40	290	32.0	683	12	BM942528	BM942528	UI-M-CG0P
41	279	30.8	516	29	AY413656	AY413656	Pan trogl
42	276.6	30.5	337	14	CB606296	CB606296	AMGNNUC:S
43	248	27.4	801	14	CK130134	CK130134	AGENCOURT
44	245.6	27.1	424	10	AM461438	AM461438	BP230005A
45	239.2	26.4	881	14	CF266082	CF266082	AGENCOURT

ALIGNMENTS

RESULT 1
LOCUS BQ721265
DEFINITION AGENCOURT 8294508 lupo: lymphatic_trunk Homo sapiens cDNA clone IMAGE:619445 5', mRNA sequence.
ACCESSION BQ721265
VERSION BQ721265.1 GI:21860162
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/ (bases 1 to 898)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-roman@ncl.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Plate: L1AM1359 row: 1 column: 08
High quality sequence stop: 669.

FEATURES
SOURCE

Location/Qualifiers
1. .898

2.89

```
/organism="Homo sapiens"
/mol_type="mRNA"
```

```
/mol_type="mRNA"
```

```
/dev=...  
/clone="IMAGE:6194455"  
/dev=...
```

```
/sex="male"
```

```
sex="male"
```

```

/issue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"

```

```
age="adult", 16 yr"
lab_host="Drosophila"
```

```
/clone_lib="Lupski_sympathetic_trunk"
```

```

--f-ont-sympacnetic trunk
/note="Vector: PCMV-SPORT6 (1.11e Tech

```

Directionally cloned using the following primers.

following adaptors:

5'-TCGACCCAGCGCTCCG-3' and
5'-GACTTACTGGGATGTTG-3'

kb for average insert size

Library, non-amplified length 1.9 kb. This is a primary

Technologies and donated by Life

College of Medicine); available through Lupski, M.D./Ph.D. (Baylor

...through life technologies."

ORIGIN

Query Match	88.3%;	Score 799.8;	DB 13;	Length 898;
Best Local Similarity	96.8%;	Pred. No. 6.5e-228;		
Matches 859; Conservative	0;	Mismatches 22;	Indels	

Oy 8 ATGACGATGTTGAACTCAAAATTGTACTGAGAGTTGGTGGCTCTGCTGGAAAGTGGAA 67
 Db 12 ATGAGATGATGGAACTCAAAATTGTACTGAGAGTTGGTGGCTCTGCGAAAGTGGAA 71
 Oy 68 CGAAAAAAGTGTGCTAATGTAAGCGGGCCATTTGTGAAATACAAATACATCCCACTT 127
 Db 72 CGGAAAAAGTGTGCTAATGTAATGCGGGCCATTTGTGAAATACAAATACATCCCACTT 131
 Oy 128 TGAAGGCATTAATATCACTGCTCCAAAGCTTATGAAGGGAAGTTGGAAACAGACAAAG 187
 Db 132 TGGAAAGCATTAATATCACTGCTCCAAAGCTTATGAAGGGAAGTTGGAAACAGACAAAG 191
 Oy 188 TGTTAATTTACAGAGCTTATCAGACATTCAGCGGAAACATTAAGTTGACATTTGATGCAGTC 247
 Db 192 TGTTAATTTACAGAGCTTATCAGACATTCAGCGGAAACATTAAGTTGACATTTGATGCAGTC 251
 Oy 248 AGAAGTTGAGTTTACGATCAAAAGTCCCAAGATGTTGCTCTCTCTCTTCTGAGACGT 307
 Db 252 AGAAGTTGAGTTTACGATCAAAAGTCCCAAGATGTTGCTCTCTCTCTTCTGAGACGT 311
 Oy 308 TTCTCATCTGATCTCTGGGTAAACTGGAAGAGCTTCAACTGTTCACTGGCTTGGCAG 367
 Db 312 TTCTCATCTGATCTCTGGGTAAACTGGAAGAGCTTCAACTGTTCACTGGCTTGGCAG 371
 Oy 368 GTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGGCTCTGTGAAGAAATCCACTTAC 427
 Db 372 GTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGGCTCTGTGAAGAAATCCACTTAC 431
 Oy 428 TCCCTACCTGCAATTTCTAGGCTTGTACTCTGTGTCACAATGCGGCCAACCGGAATTC 487
 Db 432 TCCCTACCTGCAATTTCTAGGCTTGTACTCTGTGTCACAATTGGGGCAACCGGAATTC 491
 Oy 488 TTCCCAATCTTAACTTGTGCTGCGACAGAGATGTCTCAACAAGAGAGCTGATACAGAGA 547
 Db 492 TTCCCAATCTTAACTTGTGCTGCGACAGAGATGTCTCAACAAGAGAGCTGATACAGAGA 551
 Oy 548 ATGGGATTTGTTATGTGTTAAATGACAGGTAATACCTGTCCAAAGCTGACTTATATCCCG 607
 Db 552 ATGGGATTTGTTATGTGTTAAATGACAGGTAATACCTGTCCAAAGCTGACTTATATCCCG 611
 Oy 608 AGTCTCAATTTCCGAGTGGCCGTGGAATGACAGTTTGTGAAAAATTTTGGCCGGGT 667
 Db 612 AGTCTCAATTTCCGAGTGGCCGTGGAATGACAGTTTGTGAG-AAATTTGCCGTGAT 670
 Oy 668 TGGACAAATCAGTAATTCATTTGAGAAAGCAAAAGCTCCCAATGGAATGTTCTTACATGC 727

FEATURES

Source

836

```

/OriginalBill="Homo sapiens"
/mol_type="mRNA"
/dh_mol_type="mRNA"

```

```

<!-- type="mRNA"
/db xref="taxon:9606"

```

```
/clone="IMAGE:5176724"
/lab host="DH10B"
```

```
/lab_host="DH10B"
```

```
/clone_lib="NIH_MGC_115"
```

Organ: pooled brain, lung, testis; Vector: pcMV-SPORT6; Site_1: NCI; Site_2: EGOV (destroyed); Source: [NCI](#)

male brains, age range 23-27: 1

oligo-dT primed and divested. Library is

destroyed upon cloning) (EcoRV site is

insert size range 1-3 kb. average insert size 1.8 kb, library is normalized

enriched for full-length clones and was

Research Genetics tracking code

ORIGIN

Query Match	73.4%;	Score 665.2;	DB 12;	Length 836;
Best Local Similarity	98.0%;	Pred. No. 1.3e-187;		
Matches 747;	Conservation			

	Conservative	0	Mismatches	8	Indels	7	Gaps	7
QY	147	CTGCTCCCAAGTTATGAAAGCAGAGTTGACAGGCAAAAGTTAAATTAACAGCTCAT	206					
Db	1	CTGCTCCAACTTATGAGCGAAGCTTGCAACAGGCAAAAGTTAAATTAACAGGCTCAT	60					
QY	207	CCAGCATTGAGCAAACTATAGGTGCACTTGAATTCAGTACAGAAAGTTGAGTTTACGA	266					
Db	61	CCAGCATTGAGCAAACTATAGGTGCACTTGAATTCAGTACAGAAAGTTGAGTTTACGA	119					

Db 53 GCTTACATCATGAAAGATGACATGCTTGTAGTGAAGCTTACAGATTGT 1

RESULT 4
CF72177
LOCUS
DEFINITION
CF72177 715 bp mRNA linear EST 09-OCT-2003
IMAGE:30548096 5', mRNA sequence.
CF72177
VERSION
CF72177.1 GI:37601345
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouse1.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES
source
Seq primer: PYX-5
Location/Qualifiers
1..715
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30548096"
/tissue="IMAG:30548096"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1ib="NIH BMAP HB0"
/note="Organ: Eye; Vector: PYX-Asc; Site 1: Ecor I;
Site 2: Not I; The library was constructed according
to Bonaudo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to RNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into PYX-Asc vector. The library tag
is TTAATGAAGT. This library was created for the polya tail
Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN

Query Match 62.5%; Score 565.8; DB 14; Length 715;
Best Local Similarity 92.0%; Pred. No. 7.5e-158;
Matches 597; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

258 AGTTACGATCAAGCTCCCAAGATGTCCTCTCTTCACTGTTTTCACAGT 317
Db 1 AGTTAAGATCAAGTTCCTCAAGATGTTGTTCTGTCTGTCACTCTTCTCACT 60
Qy 318 ACTTCTGGGTAACCTGAGAGAGCTTCAACTCTGTACACCGCTTGAAGTGGGTTTC 377
Db 61 ACTTCTGGGTAACCTGAGAGAGCTTCAACTCTGTACACCGCTTGAAGTGGGTTTC 377
Qy 378 TGAATCTCTGCTGTTTCCCTGGCTCTGTGAAGAAATCACTAAGTCCCTACCTG 437
Db 121 TGAATCTCTGCTGTTTCCCTGGCTCTGTGAAGAAATCACTAAGTCCCTACCTG 180

Qy 438 CATTCTCAGCCTTGTCTTACCTGTGCGCAACATGGGCGCAACCGAATTCTTCCCAATCT 497
Db 181 CATATCTCAGCCTTGTCTTACCTGTGCGCAACATGGGCGCAACCGAATTCTTCCCAATCT 240
Qy 498 TTAATCTGAGCTGCGCAGCGAGTGTCTTCAACAGAGCTGTATACAGAGATGGATTTGG 557
Db 241 CTATCTTGGCTGCGCAGCGAGTGTCTTCAACAGAGCTGTATACAGAGATGGATTTGG 300
Qy 558 TTAATCTGAGCTGCGCAGCGAGTGTCTTCAACAGAGCTGTATACAGAGATGGATTTGG 300
Db 301 CTATCTTGGCTGCGCAGCGAGTGTCTTCAACAGAGCTGTATACAGAGATGGATTTGG 360
Qy 618 CTTGCGTGTGCTGGAATGATGACCTTTTGTGAAAATTTTGGCGTGTGGAACAATC 677
Db 361 CTTGCGTGTGCTGGAATGATGACCTTTTGTGAAAATTTTGGCGTGTGGAACAATC 420
Qy 678 AGTATATTCATTGAAGAAAGCGCTTCAATGATGATGTTTGTAGTCACTGTTTNGC 737
Db 421 TGTGATTTCAATGAGAAAGCGCTTCAATGATGATGTTTGTAGTCACTGTTTNGC 737
Qy 738 TGGATCTCCCGCTCGCGCCACCAATGCTATGCTTATCATGAGAGATGACATGTC 797
Db 481 TGGATCTCTCGCTCGCGCCACCAATGCTATGCTTATCATGAGAGATGACATGTC 540
Qy 798 TTAATGATGAGCTTCAAGATTTTGTGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 857
Db 541 TTAATGATGAGCTTCAAGATTTTGTGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 600
Qy 858 TTTTCTGGGCGCAACTCTGGAATGATGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 906
Db 601 TTTTCTGGGCGCAACTCTGGAATGATGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 649

RESULT 5
CF742387
LOCUS
DEFINITION
CF742387 639 bp mRNA linear EST 10-OCT-2003
IMAGE:30619363 5', mRNA sequence.
CF742387
VERSION
CF742387.1 GI:37638726
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouse1.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES

source

Location/Qualifiers
1..639
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30619363"
/tissue="IMAG:30619363"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1ib="NIH BMAP HB0"
/note="Organ: Eye; Vector: PYX-Asc; Site 1: Ecor I;

Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Abs vector. The library tag sequence located between the Not I site and the polyA tail is TTAATGAAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Query Match 61.8%; Score 559.6; DB 14; Length 639;
Best Local Similarity 92.4%; Pred. No. 5.1e-156;
Matches 589; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```

QY 50 CTCTGCTGGAAAGTGAACGAAAGTGTCTTAATGATACCGCCATTGTGAAT 109
DB 1 CTCTGCTGGAAAGTGAACGAAAGTGTCTTAATGATACCGCCATTGTGAAT 60
QY 110 ACAATACATCCCATTTTGGAAAGCCATTATATCACTGCTCCAAAGCTATGAAGGAA 169
DB 61 ACAATACATCCCATTTTGGAAAGCCATTATATCACTGCTCCAAAGCTATGAAGGAA 120
QY 170 GGTGGCAACGACAAAGTGTATTAACAGAGCTCATCGACATTCAGCAACATTAAG 229
DB 121 GGTGGCAACGACAAAGTGTATTAACAGAGCTCATCGACATTCAGCAACATTAAG 180
QY 230 TTACACATGATTCAGACGACAAAGTGTATTAACAGATCAAGCTCCCAAGATGTGGCT 289
DB 181 TTACACATGATTCAGACGACAAAGTGTATTAACAGATCAAGCTCCCAAGATGTGGCT 240
QY 290 CTCTCTCTTGAAGCTGTTTCTCACTGTAATCTTGAGTAACGAGAAAGCTTCACT 349
DB 241 CTCTCTCTTGAAGCTGTTTCTCACTGTAATCTTGAGTAACGAGAAAGCTTCACT 300
QY 350 CTGTTCACCTGCTGACGAGTGGTGTGCTGAGTTCCTGCTGCTGCTGCTGCTG 409
DB 301 CTGTTCACCTGCTGACGAGTGGTGTGCTGAGTTCCTGCTGCTGCTGCTGCTGCTG 360
QY 410 AAGGAAATTCATCTAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 469
DB 361 AAGGAAATTCATCTAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 470 TTGGGCAACCGCAATTTCTCCCAATCTTATCTTGAGTCCGACGAGATGCTCAACA 529
DB 421 TTGGGCAACCGCAATTTCTCCCAATCTTATCTTGAGTCCGACGAGATGCTCAACA 480
QY 530 AAGAGCTGATACGACGAAATGGATGGTATGTGTTAAATGCGAGCTATACCTGCCAA 589
DB 481 AAGAGCTGATACGACGAAATGGATGGTATGTGTTAAATGCGAGCTATACCTGCCAA 540
QY 590 AGCTGACCTTATCCCGAGTCTCATTTCTGCGGTGCTGCTGAGTGAAGCTTTTGG 649
DB 541 AGCTGACCTTATCCCGAGTCTCATTTCTGCGGTGCTGCTGAGTGAAGCTTTTGG 600
QY 650 AGAAATTTTGGCGGTGGAACAATCAAGTATTTCA 688
DB 601 AGAAATTTTGGCGGTGGAACAATCAAGTATTTCA 639

```

RESULT 6
CF532917
LOCUS
DEFINITION
CF532917
ACCESSION
CF532917.1
KEYWORDS
SOURCE

656 bp mRNA linear EST 12-SEP-2003
UI-M-GH0-cgw-n-02-0-UI-r1 NIH_BMAP_GH0 Mus musculus cDNA clone
IMAGE:3057529 5', mRNA sequence.
GI:34584885
Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 656)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES

Seq primer: pX-5.
Location/Qualifiers
1..656
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:3057529"
/tissue_type="whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="PH108 (T1 phage resistant)"
/clone="NIH_BMAP_GH0"
/note="Organ: Brain; Vector: pX-Abs; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Abs vector. The library tag sequence located between the Not I site and the polyA tail is CGAACTGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 61.6%; Score 558.2; DB 14; Length 656;
Best Local Similarity 92.4%; Pred. No. 1.4e-155;
Matches 598; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

```

QY 44 TGGTGGCTCTGCTGGAAGTGAACGAAAGTGTCTTAATGATACCGGCCA-TTT 102
DB 1 TGGTGGCTCTGCTGGAAGTGAACGAAAGTGTCTTAATGATGAGCGCA-TTTT 60
QY 103 GTGGAATACATACATCCCAATTTTGAAGCATTATATCAAGCTCCAACTTATG 162
DB 61 GTGGAATACATACATGCTCTCACTTTTGAAGCATTATATCAAGCTCCAACTGATG 120
QY 163 AAGCGAAGTTGCAACAGACAAAGTGTATTAACAGCTCATCGACATTCAGCGAAA 222
DB 121 AAGCGAAGTTGCAACAGACAAAGTGTATTAACAGCTCATCGACATTCAGCGAAA 180
QY 223 CATAGGTTGACATTAATGCACTGCAAGAGTTTATGATTAAGTCAAGCTCCCAAG 282
DB 181 CATAGGTTGACATTAATGCACTGCAAGAGTGTATTAATGATTAAGTCCCAAGAT 240
QY 283 GTTGCTCTCTCTCTTCAAGCTGTTTCTCACTGTAATCTTGAGTGAATCTGAGAGAGC 342
DB 241 GTTGCTCTCTCTCTTCAAGCTGTTTCTCACTGTAATCTTGAGTGAATCTGAGAGAGC 300
QY 343 TTCAACTCTGTTCACTGCTGACAGTGGTGTGCTGAGTTCCTGTTGTTCCCTGGC 402
DB 301 TTCAACTCTGTTCACTGCTGACAGTGGTGTGCTGAGTTCCTGTTGTTCCCTGGC 360

```

QY 403 CTCTGTGAAGAAAAATCCACTTACTGCTTACCTGATTTCTCAGCTTGCTTACCTGTT 462
 Db 361 CTCTGTGAAGAAAAATCCACTTACTGCTTACCTGATTTCTCAGCTTGCTTACCTGTT 420
 QY 463 GCCAATCTTGGGCGCAACCGCAATTTCTTCCCAATCTTATCTTGGCTGCGCAGAGATGTC 420
 Db 421 GCGAATCTTGGGCGCAACCGCAATTTCTTCCCAATCTTATCTTGGCTGCGCAGAGATGTC 480
 QY 523 CTCACAGAAGAGCTGATACAGACAAATGGGATTTGTTATGTTAAATGCCAGCTTATCC 582
 Db 481 CTCACAGAAGAGCTGATACAGACAAATGGGATTTGTTATGTTAAATGCCAGCTTATCC 540
 QY 583 TGTCCAAAGCTGATCTTATCTTCCGAGTCTCATTTCTGCGTGTGCTGTGATATACAGC 642
 Db 541 TGTCCAAAGCTGATCTTATCTTCCGAGTCTCATTTCTGCGTGTGCTGTGATATACAGC 600
 QY 643 TTTTGTGAGAAAATTTTGGCGTGTGGACAAATCAGTATGATTTGAT 689
 Db 601 TTTTGTGAGAAAATCTCAGATGTTGACAAATCTGTGATTTGAT 647

RESULT 7
 BU704078 769 bp mRNA linear EST 15-JUL-2003
 LOCUS BU704078 IMAGE:6406486 5', mRNA sequence.
 ACCESSION BU704078
 VERSION BU704078.1 GI:23631836
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 769)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rewell.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 DNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/MLN at: http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
 Source
 Location/Qualifiers
 1. 769
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6406486"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="MDH08 (T1 phage resistant)"
 /clone_lib="NIH_BMAP_F00"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: Bcor I; Site 2: Not 1; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT size selected according to Not 1 site. Double strand cDNA was with Bcor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag is TGAGAGGCC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

ORIGIN
 Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

Query Match 60.9%; Score 551.4; DB 13; Length 769;
 Best Local Similarity 91.9%; Pred. No 1.6e-153;
 Matches 582; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 274 TCCCAAGATGTTGCTCTCTCTTCAAGCTGTTTCTCACTGATCTTGGGTAACG 333
 Db 3 TCCCAAGATGTTGCTCTCTCTGAGACGCTTTCTCACTGATCTTGGGTAACG 62
 QY 334 GAGAAAGCTTCAACTGTTCACTGCTGAGGAGGTTGCTGAGTCTCTGTTGT 393
 Db 63 GAGAAAGCTTCAACTGTTCACTGCTGAGGAGGTTGCTGAGTCTCTGTTGT 122
 QY 394 TTCCGAGCTCTGTGAGAGAAATCCACTTATGTCCTTACCTGATTTCTGAGCTTGC 182
 Db 123 TTCCGAGCTCTGTGAGAGAAATCCACTTATGTCCTTACCTGATTTCTGAGCTTGC 182
 QY 454 TTACTGTTGCCAACAATTGGGCCAACCCGAATTTCTCCCAATCTTATCTTGGCTCCAG 513
 Db 183 TTACTGTTGCCAACAATTGGGCCAACCCGAATTTCTCCCAATCTTATCTTGGCTCCAG 242
 QY 514 CGAATGTCCTCAACGAGAGCTGATACAGCAAAATGGGATTTGTTATGTTAAATGCC 573
 Db 243 CGAATGTCCTCAACGAGAGCTGATACAGCAAAATGGGATTTGTTATGTTAAATGCC 302
 QY 574 AGCTATCTGTGCCAAGCTGATTTATCCCGAGTCTATTTCCGCTGCTGCTG 633
 Db 303 AGCAATACCTGTGCCAAGCTGATTTATCCCGAGTCTATTTCCGCTGCTGCTG 362
 QY 634 AATGACAGCTTTTGGAGAAAATTTGCCGTTGTGACAAATCAGTATGATTTGATGAG 693
 Db 363 AATGACAGCTTTTGGAGAAAATTTGCCGTTGTGACAAATCAGTATGATTTGATGAG 422
 QY 694 AAGGAAAGCTTCCAAATGATGTTCTAGTCACTGTTATGCTGGGATCTCCGCTCC 753
 Db 423 AAGGAAAGCTTCCAAATGATGTTCTAGTCACTGTTATGCTGGGATCTCTGCTCC 482
 QY 754 GCCAATCTGCTATCGCTCATCATCAGAGAGATGAGATCTTATAGATAGCTTAC 813
 Db 483 GCCAATCTGCTATCGCTCATCATCAGAGAGATGAGATCTTATAGATAGCTTAC 542
 QY 814 AGATTGTGAAGAAAAAGACTTACTATATCTCCAACTTCAATTTTGGGGCACTC 873
 Db 543 AGATTGTGAAGAAAAAGACTTACTATATCTCCAACTTCAATTTTATTTATGAGCACTC 602
 QY 874 CTGACTATGAGAAAGATTAAGAACGAGCT 906
 Db 603 ATGACTATGAGAAAGATTAAGAACGAGCT 635

RESULT 8
 BF472046 650 bp mRNA linear EST 04-DEC-2000
 LOCUS BF472046 UI-M-BH3-awu-d-10-0-UI_r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
 DEFINITION UI-M-BH3-awu-d-10-0-UI_r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
 ACCESSION BF472046
 VERSION BF472046.1 GI:11541229
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 650)
 Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 889548

COMMENT

Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: MEST@mail.nih.gov
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
 should be noted that Bento Soares is generating a small number of
 additional specialized non-redundant arrays of BMAP cDNAs whose
 availability will be considered under appropriate and limited
 collaborative arrangements
 Seq primer: M13 Reverse

FEATURES

Source

Location/Qualifiers
 1. 650
 /organism="Mus musculus"
 /mol_type="cDNA"
 /strain="CS7BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-aw-d-10-0-UI"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_1ib="NIH BMAP M S4"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The
 NIH BMAP M S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH BMAP M S4,
 NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S3.1,
 NIH BMAP M S2, NIH BMAP M S1. The subtracted library
 (NIH BMAP M S4) was constructed as follows: PCR amplified
 cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2, and
 NIH BMAP M S3.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1
 libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH BMAP M S4 library. This procedure has been previously
 described (Bonaldi, Lennon and Soares, Genome Research
 6:791-806, 1996)"

ORIGIN

Query Match 60.1%; Score 544.8; DB 10; Length 650;
 Best Local Similarity 93.1%; Pred. No. 1,4e-151;
 Matches 603; Conservative 0; Mismatches 42; Indels 3; Gaps 3;

11 AGATGATGGAACTCAATTTGACTGAGAGGTTGGTGTCTGCTGGAAGTGAACGG 70
 Db 1 AGATGATGGAACTCAATTTGACTGAGAGGTTGGTGTCTGCTGGAAGTGAACGG 60
 Qy AAAAGGCGCTGAATGATGAGCGCGCATTTGTGAATCAATATCCCAATTTTGG 130
 Db 61 AAAAA-TCGCGCTGAATGATGAGCGCGCATTTGTGAATCAATATCCCAATTTTGG 119
 Qy 131 AAGCCATTAATCACTGCTCAAGCTTATGAAGCGAAGTTGCAACAGCAAAAGTGT 190
 Db 120 AAGCCATTAATCACTGCTCAAGCTTATGAAGCGAAGTTGCAACAGCAAAAGTGT 179
 Qy 191 TAATTACAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATTGATGAGTCAGA 250
 Db 180 TAATTACAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATTGATGAGTCAGA 239
 Qy 251 AGGTTGAGTTAGATCAAGATCCCAAGATGTGTGCTCTCTCTGAGACGTGTTTC 310

Db 240 AGGTGGTAGTTATGATCAAAAGTCCCAAGATTTGGTTCTCTGTGTCAGACTGCTTC 299
 Qy 311 TCACGTACTTCTGGGTAAACCTGAGAAAGCTTCAACTCTGTTCACTGCTGCAGGTG 370
 Db 300 TCACGTACTTCTGGGTAAACCTGAGAAAGCTTCAACTCTGTTCACTGCTGCAGGTG 359
 Qy 371 GGTTCGAGTCTCTGCTGTTTCCCTGCGGCTCTGTGAAGAAATCCACTGTGTC 430
 Db 360 GCTTTCGAGTCTCTGCTGTTTCCCTGCGGCTCTGTGAAGG-AAATCCACTGTGTC 418
 Qy 431 CTACCTGATTTCTCAGCCTTGTCTTACCTGTTGCCAATTTGGCCCAATCTTC 490
 Db 419 CTACCTGATTTCTCAGCCTTGTCTTACCTGTTGCCAATTTGGCCCAATCTTC 478
 Qy 491 CCAATCTTATCTTGGCTGCGCAGAGATGCTCTCAACAGAGCTGATACAGAGATG 550
 Db 479 CCAATCTTATCTTGGCTGCGCAGAGATGCTCTCAACAGAGCTGATACAGAGATG 538
 Qy 551 GCAATGCTTATGTTTAAATGCGAGCTATACCTGCTCAAGCCTGATCCCGAGT 610
 Db 539 GCAATGCTTATGTTTAAATGCGAGATACCTGCTCAAGCCTGATCCCGAGT 598
 Qy 611 CTCAATTCCTGCGCTGCTGCTGTAATGACAGCTTTTGTGAGAAATTT 658
 Db 599 CTCAATTCCTGCGAGTCTGTAATGAC-CTTTGTGAGAAATCT 645

RESULT 9
 AM847425/c
 LOCUS 602 bp mRNA linear EST 19-MAY-2000
 DEFINITION RCL-CT0206-270999-021-g01 CT0206 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM847425
 VERSION AM847425.1 GI:7942942
 KEYWORDS EST.

SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 1 (bases 1 to 602)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bahia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL
 MEDLINE
 PUBLISHED
 20020663
 10737800

COMMENT
 Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scr/ps/gethtml2.pl?cl=est-RCL-CT0206-270
 999-021-g01&ct3=1999-09-27&cl=1)
 Seq primer: puc 18 forward
 High quality sequence start: 8
 High quality sequence stop: 602.
 Location/Qualifiers

FEATURES

Source

1. 602
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_1ib="CT0206"

/note="Organ: Colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

```
Query Match      59.8%; Score 541.4; DB 10; Length 602;
Best Local Similarity 98.0%; Pred. No. 1.4e-150;
Matches 580; Conservative 0; Mismatches 7; Indels 5; Gaps 3;

QY 169 AGCTTGCAACGACGAAAGTGTAAATACAGAGCTATCCAGCATTCAGCAATACATTAAG 228
DB 595 AGCGTGCAACGACGAAAGTGTAAATACAGAGCTATCCAGCATTCAGCAATACATTAAG 536
QY 229 GTTACATTTGATTCAGTCAAGAGTGTAAATACAGATCAAAAGCTCCCAAGATGTGCC 288
DB 535 GTTACATTTGATTCAGTCAAGAGTGTAAATACAGATCAAAAGCTCCCAAGATGTGCC 477
QY 289 TCTCTCTTCACAGCTGTTTCTCTCACTGTACTGTCTGGTAAACTGAGAAAGCTTCAAC 348
DB 476 -CTCTCTTCACAGCTGTTTCTCTCACTGTACTGTGGTAAACTGAGAAAGCTTCAAC 418
QY 349 TCTGTTCACCTGCTGCAAGTGGTGTGGTGTCTGCTGTCTTCCCTGGCTCTGT 408
DB 417 TCTGTTCACCTGCTGCAAGTGGTGTGGTGTCTGCTGTCTTCCCTGGCTCTGT 358
QY 409 GAAGAAATTCACCTGTACTGCTCTTCACTGTACTGTCTGCTGTCTTCCCTGGCTCTGT 358
DB 357 GAAGAAATTCACCTGTACTGCTCTTCACTGTACTGTCTGCTGTCTTCCCTGGCTCTGT 298
QY 469 ATTGGGCCAACCCGAAATCTTCCCAATCTTATCTGGTCCGACGAGATGTCTCAAC 528
DB 297 ATTGGGCCAACCCGAAATCTTCCCAATCTTATCTGGTCCGACGAGATGTCTCAAC 238
QY 539 AAGAGCTGATACGACGAAATGGATGTGTATGTGTAAATCCAGCTATACCTGTCAAC 588
DB 237 AAGAGCTGATACGACGAAATGGATGTGTATGTGTAAATCCAGCTATACCTGTCAAC 178
QY 589 AAGCTGACCTTATCCCGAGTCTCAATTTCTGCGGTGCTCTGTAATGACAGCTTTTGT 648
DB 177 AAGCTGACCTTATCCCGAGTCTCAATTTCTGCGGTGCTCTGTAATGACAGCTTTTGT 118
QY 649 GAGAAATTTTGCCTGGTGTGCAATCAATCAATGATTTCAATTGAGAAAGAGCTTCC 708
DB 117 GAGAAATTTTGCCTGGTGTGCAATCAATCAATGATTTCAATTGAGAAAGAGCTTCC 708
QY 709 AATGATGTTTGTAGTGCATGTTTGTAGCTGGATCTCCCGCTCCGACCA 760
DB 60 CATGATGTTTGTAGTGCATGTTTGTAGCTGGATCTCCCGCTCCGACCA 9
```

RESULT 10
AM847426/c 595 bp mRNA linear EST 19-MAY-2000
LOCUS RCI-CT0206-270999-021-g06 CT0206 Homo sapiens cDNA, mRNA sequence.
DEFINITION AM847426
VERSION AM847426.1 GI:7942943
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL
MEDLINE
PUBMED
COMMENT
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=kt2=RCI-CT0206-270
999-021-g06&cl3=1999-09-27&cl4=1)
Seq primer: puc 18 forward
High quality sequence stop: 595.
Location/Qualifiers
1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0206"
/note="Organ: Colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source

ORIGIN

```
Query Match      59.1%; Score 535.6; DB 10; Length 595;
Best Local Similarity 97.6%; Pred. No. 7.6e-149;
Matches 576; Conservative 0; Mismatches 9; Indels 5; Gaps 3;

QY 169 AGCTTGCAACGACGAAAGTGTAAATACAGAGCTATCCAGCATTCAGCAATACATTAAG 228
DB 585 AGCGTGCAACGACGAAAGTGTAAATACAGAGCTATCCAGCATTCAGCAATACATTAAG 526
QY 229 GTTACATTTGATTCAGTCAAGAGTGTAAATACAGATCAAAAGCTCCCAAGATGTGCC 288
DB 525 GTTACATTTGATTCAGTCAAGAGTGTAAATACAGATCAAAAGCTCCCAAGATGTGCC 467
QY 289 TCTCTCTTCACAGCTGTTTCTCTCACTGTACTGTCTGGTAAACTGAGAAAGCTTCAAC 348
DB 466 -CTCTCTTCACAGCTGTTTCTCTCACTGTACTGTGGTAAACTGAGAAAGCTTCAAC 408
QY 349 TCTGTTCACCTGCTGCAAGTGGTGTGGTGTCTGCTGTCTTCCCTGGCTCTGT 408
DB 407 TCTGTTCACCTGCTGCAAGTGGTGTGGTGTCTGCTGTCTTCCCTGGCTCTGT 348
QY 409 GAAGAAATTCACCTGTACTGCTCTTCACTGTACTGTCTGCTGTCTTCCCTGGCTCTGT 348
DB 407 TCTGTTCACCTGCTGCAAGTGGTGTGGTGTCTGCTGTCTTCCCTGGCTCTGT 348
QY 409 GAAGAAATTCACCTGTACTGCTCTTCACTGTACTGTCTGCTGTCTTCCCTGGCTCTGT 348
DB 409 GAAGAAATTCACCTGTACTGCTCTTCACTGTACTGTCTGCTGTCTTCCCTGGCTCTGT 348
QY 347 GAAGAAATTCACCTGTACTGCTCTTCACTGTACTGTCTGCTGTCTTCCCTGGCTCTGT 348
DB 347 GAAGAAATTCACCTGTACTGCTCTTCACTGTACTGTCTGCTGTCTTCCCTGGCTCTGT 348
QY 469 ATTGGGCCAACCCGAAATCTTCCCAATCTTATCTGGTCCGACGAGATGTCTCAAC 528
DB 287 ATTGGGCCAACCCGAAATCTTCCCAATCTTATCTGGTCCGACGAGATGTCTCAAC 228
QY 539 AAGAGCTGATACGACGAAATGGATGTGTATGTGTAAATCCAGCTATACCTGTCAAC 588
DB 227 AAGAGCTGATACGACGAAATGGATGTGTATGTGTAAATCCAGCTATACCTGTCAAC 588
QY 227 AAGAGCTGATACGACGAAATGGATGTGTATGTGTAAATCCAGCTATACCTGTCAAC 588
DB 227 AAGAGCTGATACGACGAAATGGATGTGTATGTGTAAATCCAGCTATACCTGTCAAC 588
QY 589 AAGCTGACCTTATCCCGAGTCTCAATTTCTGCGGTGCTCTGTAATGACAGCTTTTGT 648
DB 167 AAGCTGACCTTATCCCGAGTCTCAATTTCTGCGGTGCTCTGTAATGACAGCTTTTGT 108
QY 649 GAGAAATTTTGCCTGGTGTGCAATCAATCAATGATTTCAATTGAGAAAGAGCTTCC 708
DB 107 GAGAAATTTTGCCTGGTGTGCAATCAATCAATGATTTCAATTGAGAAAGAGCTTCC 708
```


PUBMED 11076861
4
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3325)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirotsune, M., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Sano, H., Sasaki, D., Shibata, K., Sakai, K., Sakai, K., Sakazume, N., Segabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaki, T., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, Fax: 81-45-503-9216)
COMMENT
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
location/Qualifiers
1. 3325
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:9530081F05"
/db_xref="MG:2399973"
/db_xref="taxon:10090"
/clone="9530081F05"
/sex="male"
/issue_type="urinary bladder"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
460. 1080
/note="unnamed protein product; PRUNCATD MAPK PHOSPHATASE 7 homolog (Homo sapiens) (SPTR) [Q96Q82, evidence: EASYT, putative]"
/codon_start=1
/protein_id="BAC29138.1"
/db_xref="GI:26330816"
/translation="MAHMICTQIVTESLVALNLSGSEKVLIDSPREYNTSHILE AININCKLMKRLQODKVLITELIQSAKHYDIDNOKVYVYVDDSDVGLSSDC FLTVLGLTERSFVNLALAGFAEPSPRCPCERKSTLVFPCISQPCPLPVANIGPT RILPNLYICQDVLNKKVASKSLISLVSASKAIVVHYSWNTL"

ORIGIN

Query Match 53.3%; Score 482.8; DB 11; Length 3325;
Best Local Similarity 94.0%; Pred. No. 1.3e-132;
Matches 502; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
1 ATGCCCATGATGATGGAGTCAATTTGTTACTGAGAGGTGGTGGCTCTGCTGGA 60

Db
460 ATGCCCATGATGATGATGGAGTCAATTTGTTACTGAGAGGTGGTGGCTCTGCTGGA 519
Qy 61 AGTGAACCGGAAATGCTGCTAAATGATAGCCGCGCATTTGTGAATAACATACATCC 120
Db 520 AGTGAACCGGAAATGCTGCTAAATGATAGCCGCGCATTTGTGAATAACATACATCC 579
Qy 121 CACATTTTGAAGCCATTATATATCACTGCTCAAGCTTATGAAGCCGAGTTGCAAG 180
Db 580 CACATTTTGAAGCCATTATATATCACTGCTCAAGCTTATGAAGCCGAGTTGCAAG 639
Qy 181 GACAAAGTTTATATACAGAGCTATCCAGCATTCAGCGAAGATAGTTGATTTGAT 240
Db 640 GACAAAGTTTATATACAGAGCTATCCAGCATTCAGCGAAGATAGTTGATTTGAT 699
Qy 241 TCGAGTCGAAAGTTGATAGTTTATGATCAAGAGCTCCCAAGATTTGCTCTCTTCA 300
Db 700 TCGAGTCGAAAGTTGATAGTTTATGATCAAGAGCTCCCAAGATTTGCTCTCTTCA 759
Qy 301 GACTGTTTCTACGTATCTTCTGGTAACTGGAAGAGCTTCAACTCTGTCACCTG 360
Db 760 GACTGTTTCTACGTATCTTCTGGTAACTGGAAGAGCTTCAACTCTGTCACCTG 819
Qy 361 CTTCAGAGTGGTGTGCTGAGTCTCTCGTGTTCCTGGGCTCTGGAAGAAATTC 420
Db 820 CTTCAGAGTGGTGTGCTGAGTCTCTCGTGTTCCTGGGCTCTGGAAGAAATTC 879
Qy 421 ACTTAGTCCCTACCTGCAATTTCTGAGCTTCTTATCTGTTGCCAATTTGGCCAAAC 480
Db 880 ACTTAGTCCCTACCTGCAATTTCTGAGCTTCTTATCTGTTGCCAATTTGGCCAAAC 939
Qy 481 CGAATTTCTCCAACTTTATCTGAGCTGCGCAGCAGATGCTTCAACAAGAG 534
Db 940 CGAATTTCTCCAACTTTATCTGAGCTGCGCAGCAGATGCTTCAACAAGAG 993

RESULT 13
BU946569/c 792 bp mRNA linear EST 17-OCT-2003
LOCUS 70376991.001 RALUTX1.01 Rattus norvegicus cDNA, mRNA sequence.
DEFINITION BU946569
ACCESSION BU946569.1 GI:37701886
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 792)
AUTHORS
TITLE
JOURNAL
REFERENCE
FANTOM
COMMENT
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel.: 6508454102
Email: gfu@incyte.com.

FEATURES
source
1. 792
location/Qualifiers

/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/issue_type="lung"
/clone_id="RALUTX1.01"
/note="Rat, Lung, Clofibrate, 250 mg/kg, 12hr-day 28 sacrifice, SD, M/F, Pool"

ORIGIN

Query Match 47.8%; Score 433; DB 13; Length 792;
Best Local Similarity 86.2%; Pred. No. 4.9e-118;
Matches 525; Conservative 0; Mismatches 80; Indels 4; Gaps 4;

```

QY 242 GCAGTCAGAGGTGTGATTTACATCAAGCTCCCAAGTGTGC-CTCTCTCTTCA 300
DB 781 GCACTGAGAGGTGTGTGTATATATCAAGTTCCAGATGTGTCTCTGCTCAG 722
QY 301 GACTGTTTCTCACTGACTTCTGGGTAACTGGAGAGACTTCAACTCT-CTTCACT 359
DB 721 CCGGCTTTCTCACTGCTTCTGTGGTACCTGGAGAGAGACTTCAACTCTGCTTCACT 662
QY 360 GCTTTCAGGTGGGTTC-TCAGTCTCTGTTTCTTCTGCTGCTCTGTGAAGAAAT 418
DB 661 GCTTTCAGGTGGGTTC-TCAGTCTCTGTTTCTTCTGCTGCTCTGTGAAGAAAT 602
QY 419 CCACTGATGCTCCCTACCTGATTTCTGAGCTTCTTACCTGTTTGCACATTTGGGCAA 478
DB 601 CCACTGATGCTCCCTACCTGATTTCTGAGCTTCTTACCTGTTTGCACATTTGGGCAA 542
QY 479 CCGGATTTCTCCCAATCTTATCTTGTGCTGCGACGAGATGCTCTCAACAGAGACTGA 538
DB 541 CTGGAATTTCTCCCAATCTTATCTGCTGCGCTGCGACGAGATGCTCTCAACAGAGACTGA 482
QY 539 TACAGCAGATGGATTTGTATGTATTAATGCACTTACCTGCTCCAAAGCTTACT 598
DB 481 TGAACAGAGAGGTGTGTGTATGTCTMAATGCAACACCTGCTCCAAAGCTTACT 422
QY 599 TTAATCCCGAGTCTCATTTCTGCTGCTGCTGCTGATGACAGCTTTTGTGAAGAAAT 658
DB 421 TCAATCCCGAGTCTCATTTCTGCTGCTGCTGCTGATGACAGCTTTTGTGAAGAAAT 362
QY 659 TGCCTGCTGATGACAAATCAATGATTTTCAATTTGAGAAAGAAAGCTTCAATGATG 718
DB 361 TGCCTGCTGATGACAAATGATTTTCAATTTGAGAAAGAAAGCTTCAATGATG 303
QY 719 TTCTAGTGCCTGTTTGTGTGGATCTTCCGCTCCGACATGCTTATGCTTACATCA 778
DB 302 TGTCTGATCCACTGCTGTGGCTGTGATCTCCGCTCCGACATGCTTATGCTTATCA 243
QY 779 TGAAGAGATGACATGCTTGTATGATGATGATGATGATGATGATGATGATGATGAT 838
DB 242 TGAAGAGATGACATGCTTGTATGATGATGATGATGATGATGATGATGATGATGAT 183
QY 839 CTATATCTC 847
DB 182 CTCCTTCCC 174

RESULT 14
BI657528 946 bp mRNA linear EST 12-SEP-2001
LOCUS 603281866F1_NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5326049 5',
DEFINITION mRNA sequence.
ACCESSION BI657528
VERSION BI657528.1 GI:15571764
KEYWORDS Mus musculus (house mouse)
SOURCE EST.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 946)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Straubeberg, Ph.D.
Email: cga@rs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1AM1827 row: d column: 18
High quality sequence stop: 704.

```

```

FEATURES
    source
        Location/Qualifiers
            1..946
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="NMRI"
                /db_xref="taxon:10090"
                /clone="IMAGE:5326049"
                /tissue_type="tumor, gross tissue"
                /dev_stage="5 months"
                /lab_host="DH10B"
                /lab_host="NCI CGAP Mam4"
                /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                Library constructed by Life Technologies, Investigators
                providing samples: Lothar Hennighausen/Priscilla Furch,
                NIH Reference for transgenic model: Li et al., Cell Growth
                and Differentiation 7, 3-11 (1996)."
```

ORIGIN

```

Query Match      47.5%; Score 430.8; DB 12; Length 946;
Best Local Similarity 91.6%; Pred. No. 2.5e-117;
Matches 467; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

QY 398 CTGGCTCTGTGAGAGAAATCCACTGTAGTCCCTACCTGATTTCTCAGCTTGTCTTAC 457
DB 1 CTGGCTCTGTGAGAGAAATCCACTGTAGTCCCTACCTGATTTCTCAGCTTGTCTTAC 60
QY 458 CTGTTCGCAACATTTGGGCAACCCGAAATTTCTCCAACTTTATCTTGTGCTGCGAC-CGA 516
DB 61 CTGTTCGCAACATTTGGGCAACCTGATGCAAGAAATTTCTCCAACTTTATCTTGTGCTGCGACGA 120
QY 517 GATGCTCTGCAAGAGAGCTGATGACAGCAAGATGGATGGTATGCTTAAATGCCAGC 576
DB 121 GATGCTCTGCAAGAGAGCTGATGACAGCAAGATGGATGGTATGCTTAAATGCCAGC 180
QY 577 TATACCTGTCGCAAGAGCTGATTTATCCCGAGTCTCATTTCTGCTGCTGCTGAT 636
DB 181 TATACCTGTCGCAAGAGCTGATTTATCCCGAGTCTCATTTCTGCTGCTGCTGAT 240
QY 637 GACAGCTTTTGTGAGAAATTTTGGCGGTGTGAGCAATGATGATGATGATGATGATGAT 696
DB 241 GACAGCTTTTGTGAGAAATTTTGGCGGTGTGAGCAATGATGATGATGATGATGATGAT 300
QY 697 GCAAAAGCTTCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 756
DB 301 GCAAAAGCTTCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 757 ACCATGCTATGCGCTACATCATGAGAGATGATGATGATGATGATGATGATGATGATGAT 816
DB 361 ACTATGCTATGCGCTACATCATGAGAGATGATGATGATGATGATGATGATGATGATGAT 420
QY 817 TTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 876
DB 421 TTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 877 GACTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 966
DB 481 GACTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510

RESULT 15
BF815601/c 512 bp mRNA linear EST 13-JAN-2001
LOCUS BF815601
DEFINITION WK2-C10128-041200-009-c03 C10128 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF815601
VERSION BF815601.1 GI:12150300
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 512)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

```

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,F., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.K., Reis,U.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-27049922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&cl=2-MR2-C10128-
041200-009-c03&cl3=2000-12-04&cl4=1)
Seq primer: puc 18 forward
High quality sequence start: 20
High quality sequence stop: 233.
Location/Qualifiers

FEATURES

Source

1. .512
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="C10128"
/note="Organ: colon_ins; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 47.3%; Score 428.8; DB 10; Length 512;
Best Local Similarity 94.9%; Pred. No. 76-117;
Matches 486; Conservative 0; Mismatches 22; Indels 4; Gaps 4;

QY 201 GCTCATCCGCACTTACGCGAAACATPAAGTTGACATTGACATGACAGCAAGGTTGTAGT 260
DB 512 GCTCATCCGCACTTACGCGAAACATPAAGTTGACATTGACATGACAGCAAGGTTGTAGT 453
QY 261 TTACGATCAAAAGCTCCCAAGAGTGGCTCTCTCTCAAGACGTCTTTCACCTGACT 320
DB 452 TTACGATCAAAAGCTCCCAAGAGTGGCTCTCTCTCTCAAGACGTCTTTCACAGTACT 393
QY 321 TCTGGGTAAACTGGAAGAGAGCTTCAACTCTGTTCACTGCTTGACAGGAGGTTTGCTGA 380
DB 392 TCTGGGTAAACTGGAAGAGAGAGCTTCAACTCTGTTCACTGCTTGACAGGAGGTTTGCTGA 333
QY 361 GTTCTC-TGGTTGTTTCCCTGGCCCTCTGTGAAGGAAATTCACCTCTAGTCCCTACCTGCA 439
DB 332 GTTCTCAGTGTGTTCCTGGCCCTCTGTGAAGGAAATTCACCTCTAGTCCCTACCTGCA 273
QY 440 TTTCTCAGCTTGTCTTACCTGTGGCAACATTGGGCCAACCCGAATCTTCCCAATCTTT 499
DB 272 TTTCTCAGCTTGTCTTACCTGTGGCAACATTGGGCCAACCCGAATCTTCCCAATCTTT 213
QY 500 ATCTTGGCTGCCAGGAGAG-TGTCTCAACAAGAGAGCTGATPACAGA-GAATGGGATTGG 557
DB 212 ATCTTGGCTGCCAGGAGAGCTGCTCTCAACAAGAGAGCTGATGACAGACGAATGGGATTGG 153
QY 558 TTATGTGTTAAATGCGAGCTATACCTGTCCAAAGCTGAATTATGCC-CAAGTTCATT 616
DB 152 TTATGTGTTAAATGCGAGCAATACCTGTCCAAAGCTGAATTATGCCAGAGTTCATT 93

QY 617 TCTGCGTGTCCCTGTGTAATGACAGCTTTGTGAGAAATTTTGGCCGTGTTGACAAT 676
DB 92 TCTGCGTGTCCCTGTGTAATGACAGCTTTGTGAGAAATTTTGGCCGTGTTGACAAT 33
QY 677 CAGTAGATTGATTGAGAAAGCAAAAGCCCTCC 708
DB 32 CAGTAGATTGATTGAGAAAGCTTAAGCCTTC 1

Search completed: June 22, 2004, 02:51:56
Job time : 3588.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model1

Run on: June 21, 2004, 12:17:10 ; Search time 526.548 Seconds
(without alignments)
7309.622 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 906

Sequence: 1 atggcccatgagatgatttg.....agaagattaaagaccagact 906

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N_Geneseq_29Jan04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002as:*
7: geneseq2003as:*
8: geneseq2003bs:*
9: geneseq2003cs:*
10: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	906	100.0	5450	6	ACC60559 Polynucle
2	906	100.0	5450	6	ACC60572 Polynucle
3	902.8	99.6	2118	4	AAAF30479 Human pro
4	902.8	99.6	2732	4	AAD09492 Human SGP
5	902.8	99.6	3059	6	AA515768 CDNA sequ
6	902.8	99.6	3496	6	ABK47596 CDNA enco
7	902.8	99.6	3544	5	AA514639 Human CDN
8	902.8	99.6	3544	6	ABK49402 CDNA enco
9	902.8	99.6	3766	6	ABK14474 Human pro
10	902.8	99.6	4790	6	ABN83966 Human gen
11	902.8	99.6	5145	5	ABV20833 Human pro
12	902.8	99.6	5145	5	ABV21080 Human pro
13	902.8	99.6	5145	5	ABV26680 Human pro
14	902.8	99.6	5145	5	ABV20978 Human pro
15	902.8	99.6	5145	5	ABV21092 Human pro
16	902.8	99.6	5145	5	ABV21312 Human pro
17	902.8	99.6	5145	5	ABV21316 Human pro
18	902.8	99.6	5145	5	ABV26826 Human pro
19	902.8	99.6	5145	5	ABV27131 Human pro
20	902.8	99.6	5145	5	ABV26923 Human pro
21	902.8	99.6	5145	5	ABV27135 Human pro
22	902.8	99.6	5145	5	ABV28657 Human pro
23	902.8	99.6	5145	5	ABV22827 Human pro

24	902.8	99.6	5145	5	ABV26934 Human pro
25	901.2	99.5	2102	7	ADA53105 Human cod
26	901.2	99.5	3104	6	ABN59704 Novel hum
27	901.2	99.5	5111	6	ACC60521 Polynucle
28	899.6	99.3	2966	4	AAH99685 Human pro
29	782.8	86.4	2756	6	ACC60560 Polynucle
30	727.6	80.3	749	4	AAH06539 Human CDN
31	564.8	62.3	3332	6	ABK48378 CDNA enco
32	411.2	45.4	418	5	ABV11252 Human pro
33	410.2	45.3	461	5	ABV10981 Human pro
34	410.2	45.3	461	5	ABV32398 Human pro
35	410.2	45.3	461	5	ABV41327 Human pro
36	395.4	43.6	427	5	ABV10887 Human pro
37	365.8	40.4	2377	7	ACA64956 Human pro
38	365.8	40.4	2377	7	ACA64956 Human pro
39	363.6	40.1	2453	2	AAT66758 CDNA of c
40	363.6	40.1	2453	6	AAD41236 Murine ne
41	359.8	39.7	419	5	ABV32042 Human of c
42	351	38.7	2415	2	AAT66757 CDNA of c
43	349	38.5	422	5	ABV02083 Human pro
44	340.8	37.6	346	5	ABV11245 Human pro
45	339.8	37.5	345	5	ABV02076 Human pro

ALIGNMENTS

RESULT 1	ACCE60559	ACCE60559 standard; CDNA; 5450 BP.
ID	ACCE60559	
AC	ACCE60559	
XX	19-JUN-2003	(first entry)
XX	Polynucleotide relating to the invention SEQ ID NO: 108.	
XX	Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic; antihypertensive; cardiac; cytoskeletal; gene therapy; liver disease; proliferative disorder; renal failure; cardiovascular disorder; immunological disorder; arthritis; psoriasis; congenital heart defect; congenital heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
XX	Homo sapiens.	
XX	WO200257460-A2.	
XX	25-JUL-2002.	
XX	20-DEC-2001; 2001MO-US050459.	
XX	20-DEC-2000; 2000US-025668P.	
XX	30-MAR-2001; 2001US-0280186P.	
XX	01-MAY-2001; 2001US-0287735P.	
XX	05-JUN-2001; 2001US-0295848P.	
XX	25-JUN-2001; 2001US-0300465P.	
XX	(BRIM) BRISTOL-MYERS SQUIBB CO.	
XX	Jackson DG, Feder J, Nelson T, Minter G, Ramanathan C, Lee L, Slemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D, Kyser S, Mcatee P, Suchard S, Banas D, WPI, 2002-599721/64.	
XX	P-PADB; ABR52381.	
XX	Novel polynucleotides encoding human phosphatase polypeptides useful in the prevention or treatment of e.g. proliferative and cardiovascular disorders.	
XX	Claim 1; Fig 13; 801pp; English.	
XX	The invention relates to a novel isolated nucleic acid comprising a	

Query March	100.0%;	Score 906;	DB 6;	Length 5450;
Best Local Similarity	100.0%;	Pred. No. 7.8e-279;		
Matches 906; Conservative	0;	Mismatches	0;	Totals 9

Accession	Sequence	Position
Dd	1318 AAGAGATGACATGCTTTAGATGAGACTTACGATTTTGTGAAAGAAAAAGCCCTCT	137
Qy	841 ATATCTCCAAACTCAATTTCTGGGCAACTCCTGACATACAGAGAAATTAAGAC	900
Dd	1378 ATATCTCCAACTCAATTTCTGGGCAACTCCTGACATACAGAGAAATTAAGAC	1437
Qy	901 CAGACT 906	
Dd	1438 CAGACT 1443	

ACC60572
ID ACC60572 standard; cDNA; 5450 BP.
XX

13-00N-2003 (First entry)
XX

INVESTIGATION SEQ ID NO: 147

antiproliferative; cardiatic; cycostatic; gene therapy; liver disease; immunological disorder; arthritis; psoriasis; congenital heart defect; congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly. Homo sapiens.

W0200257460-A2.

25-JUL-2002.

2001, 2001MO-US050459.

30-MAR-2001; 2001US-0280186D

05-JUN-2001; 2001US-0295848P.
3E-TIM 3001

(BBTM) PRICING

Jackson DG, Feder J, Nelson T

Krystek S, Mcatee P, Suchard S, Finger J, Todderud CG, Bassolino D

WPI; 2002-599721/64.
P-PSDB: 2002-599721/64.

Novel

disorders, e.g. proliferative and cardiovascular disorders.

Example 7; Fig 19; 801pp; English

the invention relates to a novel isolated nucleic acid comprising a polynucleotide having a sequence as set forth in any one of the

the invention has anticipated the specification. The pol-

polynucleotide may have a use in gene therapy. The

ameliorating a medical condition, e.g., a proliferative disease.

such as congenital heart defects, immunological disorders, cardiovascular disorders, and psoriasis, and arthritis.

condition or susceptibility to a particular disease, and the method of the invention is useful for diagnosing a pathological

COMMERCIAL VALUE OF THE INVENTION

... ..

```

Oy 61 AGTGAACGAAAAAGTCTGTAATGATAGCCGGCCATTGAGTAATACATACATCC 120
Db 85 AGTGAACGAAAAAGTCTGTAATGATAGCCGGCCATTGAGTAATACATACATCC 144
Oy 121 CACATTTTGGAAAGCATTAAATATCAACTGCTCAAGCTTATGAAAGCGAAGTTGCAACAG 180
Db 145 CACATTTTGGAAAGCATTAAATATCAACTGCTCAAGCTTATGAAAGCGAAGTTGCAACAG 204
Oy 181 GACAAAGCTTAAATTAACAGAGCTATCCAGCAATTAAGGAAACATTAAGTTGACATTCAT 240
Db 205 GACAAAGCTTAAATTAACAGAGCTATCCAGCAATTAAGGAAACATTAAGTTGACATTCAT 264
Oy 241 TGCAAGTCGAAAGGTTGATTAAGATCAAGAGCTCCAGAGATGTTGACCTCTCTCTTCA 300
Db 265 TGCAAGTCGAAAGGTTGATTAAGATCAAGAGCTCCAGAGATGTTGACCTCTCTCTTCA 324
Oy 301 GACTGTTTCTCACTGTAATCTCTGGGTAAGTGAAGAAAGCTTCACTCTGTTCACTG 360
Db 325 GACTGTTTCTCACTGTAATCTCTGGGTAAGTGAAGAAAGCTTCACTCTGTTCACTG 384
Oy 361 CTGAGAGTGGGTTTGTCTGAGATTTCTCTGTTGTTCCCTGGCTCTGTAAGAAATCC 420
Db 385 CTGAGAGTGGGTTTGTCTGAGATTTCTCTGTTGTTCCCTGGCTCTGTAAGAAATCC 444
Oy 421 ACTGAGTCCCTACCTGATTTCTCAGGCTTGCTTACCTGTTGCCAATTTGGGCAACC 480
Db 445 ACTGAGTCCCTACCTGATTTCTCAGGCTTGCTTACCTGTTGCCAATTTGGGCAACC 504
Oy 481 CGAATTTCTCCCAATCTTATCTTGGCTGCGCAGCAGATGTCCTCAACAGAGCTGATA 540
Db 505 CGAATTTCTCCCAATCTTATCTTGGCTGCGCAGCAGATGTCCTCAACAGAGCTGATA 564
Oy 541 CAGCAGATGGGATTTGTTATGTTAAATGCAAGCTTACTGTTCCAAAGCTGACTTT 600
Db 565 CAGCAGATGGGATTTGTTATGTTAAATGCAAGCTTACTGTTCCAAAGCTGACTTT 624
Oy 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGTAAGTGAAGCTTTTGTGAAGAAATTTG 660
Db 625 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGTAAGTGAAGCTTTTGTGAAGAAATTTG 684
Oy 661 CCGTGTGGAACAAATCAGTATGATTTCATTTGGAAGAAAGAGCTCCCAATGATGTT 720
Db 685 CCGTGTGGAACAAATCAGTATGATTTCATTTGGAAGAAAGAGCTCCCAATGATGTT 744
Oy 721 CTAGTGCACTGTTTACCTGGGATCTCCGCTCCGACACATGCTATGCTTACATCATG 780
Db 745 CTAGTGCACTGTTTACCTGGGATCTCCGCTCCGACACATGCTATGCTTACATCATG 804
Oy 781 AAGAGATGACATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 805 AAGAGATGACATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 864
Oy 841 ATATCTCCAAATCTCAATTTCTGGGCCCACTCTGAGCTATGAGAGAGATTAAGAAC 900
Db 865 ATATCTCCAAATCTCAATTTCTGGGCCCACTCTGAGCTATGAGAGAGATTAAGAAC 924
Oy 901 CAGACT 906
Db 925 CAGACT 930

```

RESULT 4

AADO9492 standard; DNA; 2732 BP.

AADO9492;

10-SEP-2001 (first entry)

Human SGP002 phosphatase polypeptide encoding DNA
 Human SGP002 phosphatase polypeptide; phosphatase-related disease;
 immune-related disorder; ocular disease; organ transplant rejection;

infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
 metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
 Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
 attention disorder; cognition disorder; psychiatric disorder; cystostatic;
 neurological disorder; vitreous; neotrophic; cerebroprotective; therapy;
 neuroprotective; antibacterial; vulnery; tranquilliser; antiepileptic;
 hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
 antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
 MKP; mitraline; chromosome 12p11.1-p12.1; 68.

Homo sapiens.

Location/Qualifiers

Key 538..2535

FT CDS /tag= a /product= "Human SGP002 phosphatase polypeptide"

MO200146394-A2.

28-JUN-2001.

21-DEC-2000; 2000MO-US034736.

21-DEC-1999; 99US-0173255P.

28-DEC-1999; 99US-0175766P.

25-JAN-2000; 2000US-0178078P.

31-JAN-2000; 2000US-0179301P.

(SUGEN-) SUGEN INC.

Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;

Hill RJ, Flanagan P;

WPI; 2001-418058/44.

P-PSDB; AAE04834.

Novel phosphatase polypeptide useful for treating cancers, immune-related

diseases and disorders, cardiovascular disease, brain or neuronal-

associated diseases and metabolic disorders.

Claim 29; Fig 1; 186pp; English.

The present invention relates to phosphatase polypeptides, nucleotide
 sequences encoding them, as well as various products and methods useful
 for the diagnosis and treatment of various phosphatase-related diseases
 and conditions. Substance that modulates the activity of phosphatase
 polypeptide is used to treat immune-related diseases and disorders,
 cardiovascular disease, brain or neuronal-associated diseases and
 metabolic disorders, including cancers of tissues, cancers of
 haematopoietic origin, diseases of central and peripheral nervous system,
 Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
 lateral sclerosis, viral infections, infections caused by prions,
 bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual
 dysfunction, mood disorders, attention disorders, cognition disorders,
 hypotension, psychotic disorders, neurological disorders,
 dyskinesias and organ transplant rejection. The present sequence is a DNA
 encoding human SGP002 phosphatase polypeptide. This sequence is
 classified as dual specificity phosphatase (DSP) and MAP kinase
 phosphatase (MKP). SGP002 gene maps to chromosomal position 12p11.1-p12.1

Sequence 2732 BP; 710 A; 684 C; 686 G; 652 T; 0 U; 0 Other;

Query Match 99.6%; Score 902.8; DB 4; Length 2732;

Best Local Similarity 99.8%; Pred. No. 5.6e-278;

Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGGCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
 538 ATGGCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 597
 61 AGTGAACGAAAAAGTCTGTAATGATAGCCGGCCATTGAGTAATACATACATCC 120

Db 367 TGCAGTGAAGGTTGATTTAGATCAAGCTCCAGAAATGTGCTCTCTCTCA 426
 Qy 301 GACTGTTTCTCACTGTAATCTTGGGTAAGTGAAGAGCTTCAACTGTTCACCTG 360
 Db 427 GACTGTTTCTCACTGTAATCTTGGGTAAGTGAAGAGCTTCAACTGTTCACCTG 486
 Qy 361 CTTCAGAGTGGGTTTGGTGAAGTCTCTGTTGTTCCCTGGCCCTGTGAAGAAATCC 420
 Db 487 CTTCAGAGTGGGTTTGGTGAAGTCTCTGTTGTTCCCTGGCCCTGTGAAGAAATCC 546
 Qy 421 ACTCTAGTCCCTACCTGATTTCTCAGCCCTGCTTACCTGTGCAACATTTGGCCAAAC 480
 Db 547 ACTCTAGTCCCTACCTGATTTCTCAGCCCTGCTTACCTGTGCAACATTTGGCCAAAC 606
 Qy 481 CGAATCTTCCCAATCTTATCTTGGCTGCTGCAAGAGATGCTTCAACAGAGCTGATA 540
 Db 607 CGAATCTTCCCAATCTTATCTTGGCTGCTGCAAGAGATGCTTCAACAGAGCTGATA 666
 Qy 541 CAGAGATGGGATGTTATGTTATGCTTAATGCGCTATACCTGTCCAAAGCCGACTTT 600
 Db 667 CAGAGATGGGATGTTATGTTATGCTTAATGCGCTATACCTGTCCAAAGCCGACTTT 726
 Qy 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Db 727 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
 Qy 661 CCGTGTGGAACAATCATGATGATTTCTTATGAGAAAGCAAAAGCTTCAATGATGTT 720
 Db 787 CCGTGTGGAACAATCATGATGATTTCTTATGAGAAAGCAAAAGCTTCAATGATGTT 846
 Qy 721 CTATGACCTGTTATGCTGGAATCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 780
 Db 847 CTATGACCTGTTATGCTGGAATCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 906
 Qy 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGGAAGAAAGAGCTTACT 840
 Db 907 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGGAAGAAAGAGCTTACT 966
 Qy 841 ATATCTCCAAATCTCAATTTCTGCGGCAACTCTCTGACATGAGAGATTAAGAAC 900
 Db 967 ATATCTCCAAATCTCAATTTCTGCGGCAACTCTCTGACATGAGAGATTAAGAAC 1026
 Qy 901 CAGACT 906
 Db 1027 CAGACT 1032

RESULT 6
 ABK47596
 ID ABK47596 standard; cDNA, 3496 BP.
 XX
 AC ABK47596;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE cDNA encoding human dual-specificity phosphatase-3 (DSP-16) protein.
 XX
 KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KW cancer; graft-versus-host disease; allergy; metabolic disease;
 KW abnormal cell growth; abnormal cell proliferation; contract inhibition;
 KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KW intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 562..2559
 FT /tag=a
 FT /product="Human dual-specificity phosphatase-3 (DSP-16)
 FT protein"
 XX

PN WO200226997-A2.
 XX
 PD 04-APR-2002.
 XX
 XX 25-SEP-2001; 2001WO-US030124.
 XX
 XX 26-SEP-2000; 2000US-0235487P.
 XX
 PA (CEPT-) CEPTYR INC.
 PI Luche RM, Wei B;
 XX
 XX WPI; 2002-315802/35.
 DR P-PSDB; AAU79156.
 XX
 FT New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 XX
 PS Claim 7; Fig 1; 87pd; English.
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity for modulation of a proliferative response in a cell, survival
 CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present nucleic
 CC acid sequence is that of the human dual-specificity phosphatase-3 (DSP-
 CC 16) gene located on chromosome 12p. This sequence encodes the human DSP-
 CC 16 protein of the invention
 XX
 SO Sequence 3496 BP; 939 A; 854 C; 856 G; 847 T; 0 U; 0 Other;
 Query Match 99.6%; Score 902.8; DB 6; Length 3496;
 Best Local Similarity 99.6%; Pred. No. 6,5e-278;
 Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 ATGGCCCATGAGATGATGGAAGCTCAATTTGTTACTGAGAGGTTGGTCTGCTGGAA 60
 Db 562 ATGGCCCATGAGATGATGGAAGCTCAATTTGTTACTGAGAGGTTGGTCTGCTGGAA 621
 Qy 61 AGTGAACGAGAAAGAGTGTCTAATTTAGATAGCCGCAATTTGTAATCAATATATCC 120
 Db 622 AGTGAACGAGAAAGAGTGTCTAATTTAGATAGCCGCAATTTGTAATCAATATATCC 681
 Qy 121 CACATTTTGAAGCCATTAATATCAATGCTCCCAAGCTTATGAAGCGAAGTTGCCAACG 180
 Db 682 CACATTTTGAAGCCATTAATATCAATGCTCCCAAGCTTATGAAGCGAAGTTGCCAACG 741
 Qy 181 GACAAAGTGAATTAATTAAGAGCTCATCAGCAATTCAGCGAAACATAAGTTGACATTGAT 240
 Db 742 GACAAAGTGAATTAATTAAGAGCTCATCAGCAATTCAGCGAAACATAAGTTGACATTGAT 801
 Qy 241 TGCAGTCAGAGGTTGATTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 300
 Db 802 TGCAGTCAGAGGTTGATTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 861
 Qy 301 GACTGTTTCTCACTGTAATCTTGGGTAAGTGAAGAGCTTCAACTGTTCACCTG 360
 Db 862 GACTGTTTCTCACTGTAATCTTGGGTAAGTGAAGAGCTTCAACTGTTCACCTG 921
 Qy 361 CTTCAGAGTGGGTTTGGTGAAGTCTCTGTTGTTCCCTGGCCCTGTGAAGAAATCC 420
 Db 922 CTTCAGAGTGGGTTTGGTGAAGTCTCTGTTGTTCCCTGGCCCTGTGAAGAAATCC 981
 Qy 421 ACTCTAGTCCCTACCTGATTTCTCAGCCCTGCTTACCTGTGCAACATTTGGCCAAAC 480

Db	598	ACTGAAAGGAAAAAGTCTGCTTAATTGATATACGCGGCATTTTGGATATCAATATACATCC	657
OY	121	CACATTTTGGAAAGCCATTATATATCAACTGCTCCAGACTTATGAGCGAAGGTTGCAACAG	180
Db	658	CACATTTTGGAAAGCCATTATATATCAACTGCTCCAGACTTATGAGCGAAGGTTGCAACAG	717
OY	181	GACAAAGTGTAAATTAACAGAGCTATCCAGGATTCAGCGAAACATAAGGTTGACATTGAT	240
Db	718	GACAAAGTGTAAATTAACAGAGCTATCCAGGATTCAGCGAAACATAAGGTTGACATTGAT	777
OY	241	TGCACTAGAAAGTTTGTAAGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA	300
Db	778	TGCACTAGAAAGGTTGTAGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA	837
OY	301	GACTGTTTTCTCACTGTACTTCTGGGTAAACTGGAGAAGAGCTTCAACTCTTTCACCTG	360
Db	838	GACTGTTTTCTCACTGTACTTCTGGGTAAACTGGAGAAGAGCTTCAACTCTTTCACCTG	897
OY	361	CTTGCAAGGTGGGTTTGCTGAAGTTCTCTCGTTGTTTCCCTGGCCTCTGTAGAGAAATCC	420
Db	898	CTTGCAAGGTGGGTTTGCTGAAGTTCTCTCGTTGTTTCCCTGGCCTCTGTAGAGAAATCC	957
OY	421	ACTCTAGTCCCTCACTGCAATTTCTCAAGCCTTGCTCACTGTTGGCAACATATGGGCCAAC	480
Db	958	ACTCTAGTCCCTCACTGCAATTTCTCAAGCCTTGCTCACTGTTGGCAACATATGGGCCAAC	1011
OY	481	CGAATCTTCCCAATCTTTATCTTGGCTGCACGCGAGATGTCCTCAACAAAGAGCTGATA	540
Db	1018	CGAATCTTCCCAATCTTTATCTTGGCTGCACGCGAGATGTCCTCAACAAAGAGCTGATA	1077
OY	541	CAGCAGAAATGGGAATGGTATGTATGTATTAATGCAAGCTATACCTGTCCAAAGCTGACATT	600
Db	1078	CAGCAGAAATGGGAATGGTATGTATGTATTAATGCAAGCTATACCTGTCCAAAGCTGACATT	1137
OY	601	ATCCCGAGTCTCATTTCTCGTGTGCTGCTGCAATGACAGCTTTTGTGAGAAATTTTG	660
Db	1138	ATCCCGAGTCTCATTTCTCGTGTGCTGCTGCAATGACAGCTTTTGTGAGAAATTTTG	1197
OY	661	CCGTGTTGGACAAATCAGATGATTTCAATTGAGAAAGCAAAAGCCTCAATGATGTGT	720
Db	1198	CCGTGTTGGACAAATCAGATGATTTCAATTGAGAAAGCAAAAGCCTCAATGATGTGT	1257
OY	721	CTAGTGCACCTGTTTATGCTGGGATCTCCCGCTCCGCGACATGCTTATGCGCTTACATCAG	780
Db	1258	CTAGTGCACCTGTTTATGCTGGGATCTCCCGCTCCGCGACATGCTTATGCGCTTACATCAG	1317
OY	781	AAGAGGATGCAATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAACCTACT	840
Db	1318	AAGAGGATGCAATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAACCTACT	1377
OY	841	ATATCTCCAAACCTCAATTTTCTGGGGCACTCTGACATATGAGAAAGATTAAGAAC	900
Db	1378	ATATCTCCAAACCTCAATTTTCTGGGGCACTCTGACATATGAGAAAGATTAAGAAC	1437
OY	901	CAGACT 906	
Db	1438	CAGACT 1443	

RESULT 5
AAS15768
ID AAS15768 standard; cDNA; 3059 BP
XX

AC AAS15768;

DT 18-JUN-2002 (first entry)

CDNA sequence of human (d),

Human: dual

KM neuronal degeneration syndrome; Alzheimer's disease; depression;
KM schizophrenia; asthma; immune disorder; inflammatory process; arthritis
KM osteoporosis; diabetes; gene therapy; chromosome 12; 88.

XX	Homo sapiens.
OS	
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	127..2124
FT	/tag= a
FT	/product= "DUSP-10 protein"
PN	MO200177340-A1.
XX	
PD	18-OCT-2001.
PF	06-APR-2001; 2001WC-EP003966.
PR	10-APR-2000; 2000EP-00107143.
PA	(MERCK) MERCK PATENT GMBH.
PI	Duecker K;
DR	WPI; 2002-010917/01.
XX	
DR	P-PDB; AAU09946.
XX	
PT	Novel dual specificity phosphatase polypeptides useful for treating
PT	cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
PT	disease, depression, schizophrenia, asthma and immune disorders.
PS	Claim 5; Page 34-37; 43pp; English.
XX	
CC	The present invention relates to a new isolated dual specificity
CC	phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
CC	sequence that is fully defined in the specification. The invention also
CC	provides a sequence encoded by a 3059 nucleotide sequence fully defined
CC	in the specification, and a sequence having at least 95 % identity to the
CC	polypeptide, or fragments or variants of DUSP-10. The invention is useful
CC	for treating cancer e.g. leukaemia, colon carcinoma, lung cancer,
CC	prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
CC	stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
CC	schizophrenia, cardiac myopathies, asthma, immune disorders,
CC	inflammatory processes e.g. arthritis, bowel disease, type I diabetes,
CC	osteoporosis, diabetes and diabetes associated diseases. The molecules of
CC	the invention are also useful as vaccines for inducing immunological
CC	response in a mammal, in disease diagnosis and in assays for screening
CC	agonistic or antagonistic compounds. Other uses of the invention include
CC	identifying membrane bound or soluble receptors, as a diagnostic reagent,
CC	in chromosome localisation studies, and as a valuable tool in tissue
CC	expression studies. The present sequence represents cDNA of the human
CC	dual specificity phosphatase, DUSP-10. DUSP-10 is located on chromosome
CC	12
SQ	Sequence 3059 BP; 831 A; 761 C; 709 G; 758 T; 0 U; 0 Other:
Query Match	99.6%; Score 902.8; DB 6; Length 3059;
Best Local Similarity	99.8%; Pred. No. 6e-278;
Matches 904; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	1 ATGGCCCATGAGATGATTGAACCTCAATTGTTTACTAGAGAGTGGNGGCTCGCTGGAA 60
Db	127 ATGGCCCATGAGATGATTGAACCTCAATTGTTTACTAGAGAGTGGTGCTGCTGGAA 186
QY	61 AGTGAAGAAGAAAAAGTCCTCTAATTGATGATGCCGCACATTGTGATAACAATACATCC 120
Db	187 AGTGAAGAAGAAAAAGTCCTCTAATTGATGATGCCGCACATTGTGATAACAATACATCC 246
QY	121 CAcATTtTGAAGcCATTAtATATCACTGCTCAAGCTTATGAAGGAAGGTGGAAAG 180
Db	247 CAcATTtTGAAGcCATTAtATATCACTGCTCAAGCTTATGAAGGAAGGTGGAAAG 306
QY	181 GAcAAAGtTAATTAAGAAGCTCACAGCAATTCAGcAAAcATMAgTTGACATtTGAAT 240
Db	307 GAcAAAGtTAATTAAGAAGCTCACAGCAATTCAGcAAAcATMAgTTGACATtTGAAT 366
QY	241 TGcAGTcGAAGgTtTAGTtTACAGATCCAAGGtcCAAAGtTtGcCTctCTCTtTCA 300

QY 601 ATCCCGAGTCTCATTTCCGCGTGTGCTGTAATGACAGCTTTTGAGAAATTTTG 660
 Db 1189 ATCCCGAGTCTCATTTCCGCGTGTGCTGTAATGACAGCTTTTGAGAAATTTTG 1248
 QY 661 CCGTGTGGACAATCAGTATGATTTTCATGAGAAAGCAAGCTCCATGATGTGT 720
 Db 1249 CCGTGTGGACAATCAGTATGATTTTCATGAGAAAGCAAGCTCCATGATGTGT 1308
 QY 721 CTAGTGACCTTTTACCTGGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 780
 Db 1309 CTAGTGACCTTTTACCTGGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 1368
 QY 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTGTGAGAAAGCAAGCTTACT 840
 Db 1369 AAGAGATGACATGCTTTAGATGAGCTTACAGATTGTGAGAAAGCAAGCTTACT 1428
 QY 841 ATATCTCCAACTTCAATTTTCTGGGCACTCTGCACTATGAGAAAGATTTAGAAC 900
 Db 1429 ATATCTCCAACTTCAATTTTCTGGGCACTCTGCACTATGAGAAAGATTTAGAAC 1488
 QY 901 CAGACT 906
 Db 1489 CAGACT 1494

RESULT 8

ABK49402 standard; cDNA; 3544 BP.

ABK49402;

02-JUL-2002 (first entry)

CDNA encoding human dual specificity phosphatase 21117 protein.

Human; dual specificity phosphatase 21117; erythroid-related disorder;

haematopoietic-related disorder; leukemias; autoimmune disease; anaemia;

erythrocystosis; liver-related disorder; cancer; gene; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 589..2586

/tag= a

/product= "Human dual specificity phosphatase 21117"

/note= "Specifically claimed in claim 2"

US2002034807-A1.

21-MAR-2002.

23-MAR-2001; 2001US-00816494.

24-MAR-2000; 2000US-0191858P.

(MEYE/) MEYERS R A.

Meyers RA;

WPI; 2002-351088/38.

P-PSDB; AAU79929.

New nucleic acids designated 38692 and 21117, encoding dual specificity

phosphatases for treating cell proliferation and differentiation

disorders including hematopoietic and erythroid-related disorders and

cancers.

CC polypeptide may be used to diagnose and treat haematopoietic-related
 CC disorders such as anaemias and autoimmune diseases, erythroid-related
 CC disorders such as anaemias and erythrocytosis, liver-related disorders,
 CC and cancers, particularly of the breast, colon, adipose, prostate and
 CC lung. The present nucleic acid sequence encodes the human dual
 CC specificity phosphatase 21117 protein of the invention, as described
 CC above

SO Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;

Query Match 99.8%; Score 902.8; DB 6; Length 3544;

Best Local Similarity 99.8%; Pred. No. 6.5e-278;

Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGATGATTTGAACTCAATTTGTTACTAGAGGTTGGTGGCTGTGGA 60
 Db 589 ATGGCCCATGATGATTTGAACTCAATTTGTTACTAGAGGTTGGTGGCTGTGGA 648
 QY 61 AGTGAACGAAAAAGTGTCTTAATTGATAGCCGCACTTTGTGAATCAATACATCC 120
 Db 649 AGTGAACGAAAAAGTGTCTTAATTGATAGCCGCACTTTGTGAATCAATACATCC 708
 QY 121 CACATTTGGAAGCCATTAAATCACTGCTCCAGCTTATGAAGCGAGTTGCAACAG 180
 Db 709 CACATTTGGAAGCCATTAAATCACTGCTCCAGCTTATGAAGCGAGTTGCAACAG 768
 QY 181 GACAAAGTGTAAATTAAGAGCTCATGACATTCAGCAAGCAACATAGTTGACATGAT 240
 Db 769 GACAAAGTGTAAATTAAGAGCTCATGACATTCAGCAAGCAACATAGTTGACATGAT 828
 QY 241 TGCAGTCAGAGGTTGATTTTAAAGATCAAGTCCCAAGATGTGCTCTCTTCA 300
 Db 829 TGCAGTCAGAGGTTGATTTTAAAGATCAAGTCCCAAGATGTGCTCTCTTCA 888
 QY 301 GACTGTTTCTACCTGACTTCTGGGTAACTGAGAAAGCTTCACTCTTCACTG 360
 Db 889 GACTGTTTCTACCTGACTTCTGGGTAACTGAGAAAGCTTCACTCTTCACTG 948
 QY 361 CTGGCAGTGGGTTGCTGAGTCTCTGTTGTTCCCTGAGCTCTGTAAGAAATTC 420
 Db 949 CTGGCAGTGGGTTGCTGAGTCTCTGTTGTTCCCTGAGCTCTGTAAGAAATTC 1008
 QY 421 ACTGTAGTCCCTACCTGACTTCTCAGCTTGTCTTACCTGTGCAACATTTGGCAACC 480
 Db 1009 ACTGTAGTCCCTACCTGACTTCTCAGCTTGTCTTACCTGTGCAACATTTGGCAACC 1068
 QY 481 CGAATTTTCCCAATCTTTATCTTGCTGCGCAGAGGATGCTTCAACAGAGCTATA 540
 Db 1069 CGAATTTTCCCAATCTTTATCTTGCTGCGCAGAGGATGCTTCAACAGAGCTATA 1128
 QY 541 CAGCAGAAATGGGATTTGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCTGACTTT 600
 Db 1129 CAGCAGAAATGGGATTTGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCTGACTTT 1188
 QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGAGAAATTTTG 660
 Db 1189 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGAGAAATTTTG 1248
 QY 661 CCGTGTGGACAATCAGTATGATTTTCATGAGAAAGCAAGCTCCATGATGTGT 720
 Db 1249 CCGTGTGGACAATCAGTATGATTTTCATGAGAAAGCAAGCTCCATGATGTGT 1308
 QY 721 CTAGTGACCTTTTACCTGGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 780
 Db 1309 CTAGTGACCTTTTACCTGGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 1368
 QY 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTGTGAGAAAGCAAGCTTACT 840
 Db 1369 AAGAGATGACATGCTTTAGATGAGCTTACAGATTGTGAGAAAGCAAGCTTACT 1428
 QY 841 ATATCTCCAACTTCAATTTTCTGGGCACTCTGCACTATGAGAAAGATTTAGAAC 900
 Db 1429 ATATCTCCAACTTCAATTTTCTGGGCACTCTGCACTATGAGAAAGATTTAGAAC 1488

DR P-PSDB; AAU09016.
XX
PT New human dual specificity polypeptides and nucleic acids for diagnosis
PT of disease and treatment of e.g. liver disorders.
XX
PS
XX
Claim 1; Fig 1; 143pp; English.

The invention relates to two novel human dual specificity phosphatases designated 2117 and 38692, the nucleic acids encoding them (including fragments, allelic variants, their complements or nucleic acids that hybridise to them) and antibodies raised against the proteins. The antibody is useful for detecting the presence of the polypeptide, and the nucleic acid fragments are useful for detecting the presence of the polypeptide, and the nucleic acid, for diagnosing disease. Agents (e.g. the antibody and antisense sequences) are useful for modulating the activity or expression of the polypeptide/nucleic acid, useful for the treatment of e.g. hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus, rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others listed in the specification) liver disorders, erythroid associated disorders (e.g. haemolytic anaemia) cellular proliferative or carcinomas, sarcomas, metastatic cancers. Agents modulating 2117 and 38692 are also useful for modulating the proliferation, survival, migration or differentiation of a 38692 or 2117-expressing cell. The polypeptide and nucleic acids are useful for identifying modulating agents. The present sequence encodes the dual specificity phosphatase 2117

Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;

DR P-PSDB; AAU09016.
XX
PT New human dual specificity polypeptides and nucleic acids for diagnosis
PT of disease and treatment of e.g. liver disorders.
XX
PS
XX
Claim 1; Fig 1; 143pp; English.

The invention relates to two novel human dual specificity phosphatases designated 2117 and 38692, the nucleic acids encoding them (including fragments, allelic variants, their complements or nucleic acids that hybridise to them) and antibodies raised against the proteins. The antibody is useful for detecting the presence of the polypeptide, and the nucleic acid fragments are useful for detecting the presence of the polypeptide, and the nucleic acid, for diagnosing disease. Agents (e.g. the antibody and antisense sequences) are useful for modulating the activity or expression of the polypeptide/nucleic acid, useful for the treatment of e.g. hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus, rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others listed in the specification) liver disorders, erythroid associated disorders (e.g. haemolytic anaemia) cellular proliferative or carcinomas, sarcomas, metastatic cancers. Agents modulating 2117 and 38692 are also useful for modulating the proliferation, survival, migration or differentiation of a 38692 or 2117-expressing cell. The polypeptide and nucleic acids are useful for identifying modulating agents. The present sequence encodes the dual specificity phosphatase 2117

Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;

QY 901 CAGACT 906
 DB 1489 CAGACT 1494

RESULT 9
 ABK14474
 ID ABK14474 standard; cDNA; 3766 BP.
 AC ABK14474;
 XX 08-MAY-2002 (first entry)
 XX Human protein phosphatase 7 (pp7) cDNA sequence.
 XX Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
 KW acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
 KW Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
 KW dementia; Parkinson's disease; developmental disorder; Down's syndrome;
 KW cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
 KW melanoma; myeloma sarcoma; gene; ss.
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 538..2535
 FT CDS /tag= a
 FT /product= "Protein_phosphatase_7 (pp7)"
 FT
 PN WO200210363-A2.
 XX
 XX 07-FEB-2002.
 XX
 XX 26-JUL-2001; 2001WO-US023716.
 XX
 XX 28-JUL-2000; 2000US-0221679P.
 PR 03-AUG-2000; 2000US-0223272P.
 PR 10-AUG-2000; 2000US-0224309P.
 PR 18-AUG-2000; 2000US-0226728P.
 PR 30-AUG-2000; 2000US-0229254P.
 PR 08-SEP-2000; 2000US-0231366P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;
 PI Stewart EA, Gandhi AR, Patterson C, Lee EA, Hafalia AJA, Lu DAM;
 PI Tribouley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;
 PI Wallia NK, Kearney LJ;
 XX
 XX WPI; 2002-188735/24.
 DR P-PSDB; AAU75789.
 XX
 XX New protein phosphatases, useful for diagnosing, treating or preventing
 PT immune system disorders (e.g. Crohn's disease), neurological disorders
 PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
 PT cancers).
 XX
 XX Claim 5; Page 114-115; 117pp; English.
 PS
 XX The present invention relates to a new polypeptide, a naturally occurring
 XX amino acid sequence at least 95% identical to it, a biologically active
 XX fragment of it or an immunogenic fragment of it. The polypeptides,
 XX polynucleotides, agonists and antagonists are useful for diagnosing,
 XX treating or preventing disorders associated with aberrant expression of
 XX protein phosphatases (PP), particularly immune system disorders e.g.
 XX acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
 XX asthma or Crohn's disease, neurological disorders e.g. epilepsy,
 XX Huntington's disease, dementia or Parkinson's disease, developmental
 XX disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
 XX cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
 XX or sarcoma. The present nucleic acid sequence encodes human protein
 XX phosphatase 7 (pp7) which is one of several human protein phosphatases

CC (AAU75783-AAU75792) of the invention
 XX
 SQ Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;
 Query Match 99.6%; Score 902.8; DB 6; Length 3766;
 Best Local Similarity 99.8%; Pred. No. 6.7e-278;
 Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGCCCATGAGATGATTGGAATCTCAAAATGTTACTGAGAGGTGGTGGCTCTGCTGGAA 60
 DB |||||
 QY 538 ATGCCCATGAGATGATTGGAATCTCAAAATGTTACTGAGAGGTGGTGGCTCTGCTGGAA 60
 DB |||||
 QY 61 AGTGAACGGAAGGCTGCTTAATTGATAGCGGCCCATTTGTGGAATACATATCC 120
 DB |||||
 QY 598 AGTGAACGGAAGGCTGCTTAATTGATAGCGGCCCATTTGTGGAATACATATCC 657
 DB |||||
 QY 121 CACATTTTGGAGCCATTAATCACTCTCAAGCTTATGAAGGAGGTGGCAACAG 180
 DB |||||
 QY 658 CACATTTTGGAGCCATTAATCACTCTCAAGCTTATGAAGGAGGTGGCAACAG 717
 DB |||||
 QY 181 GACAAAGTGTAAATTACAGAGCTCATCCAGCAATTCAGCGAAACATAAGGTTGACATTCAT 240
 DB |||||
 QY 718 GACAAAGTGTAAATTACAGAGCTCATCCAGCAATTCAGCGAAACATAAGGTTGACATTCAT 777
 DB |||||
 QY 241 TGCAGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCTTCA 300
 DB |||||
 QY 778 TGCAGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCTTCA 837
 DB |||||
 QY 301 GACTGTTTCTCAGTGTACTTCTGGGTAAACTGGGAGAGAGCTTCACTCTGTTCACTG 360
 DB |||||
 QY 838 GACTGTTTCTCAGTGTACTTCTGGGTAAACTGGGAGAGAGCTTCACTCTGTTCACTG 897
 DB |||||
 QY 361 CTTGCAAGTGGGTTGCTGAGTCTCTGTTGTTTCCCTGCTCTCTGGAAGGAAATCC 420
 DB |||||
 QY 898 CTTGCAAGTGGGTTGCTGAGTCTCTGTTGTTTCCCTGCTCTCTGGAAGGAAATCC 957
 DB |||||
 QY 421 ACTTAGTCCCTACTCCTGATTTCTCAGCTTGTCTTACTGTTGGCAACATTTGGCCAAAC 480
 DB |||||
 QY 958 ACTTAGTCCCTACTCCTGATTTCTCAGCTTGTCTTACTGTTGGCAACATTTGGCCAAAC 1017
 DB |||||
 QY 481 CGAATTTCTCCAAATCTTTTATCTTGGCTGCCAGAGATGTCCTCAACAAGGAGCTGATG 540
 DB |||||
 QY 1018 CGAATTTCTCCAAATCTTTATCTTGGCTGCCAGAGATGTCCTCAACAAGGAGCTGATG 1077
 DB |||||
 QY 541 CAGCAGAAATGGGATTTGGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 600
 DB |||||
 QY 1078 CAGCAGAAATGGGATTTGGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTT 1137
 DB |||||
 QY 601 ATCCCCAGTCTCATTTCTGCGTGTCCCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
 DB |||||
 QY 1138 ATCCCCAGTCTCATTTCTGCGTGTCCCTGTGAATGACAGCTTTTGTGAGAAATTTTG 1197
 DB |||||
 QY 661 CCGTGGTTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTCCAAATGGATGTT 720
 DB |||||
 QY 1198 CCGTGGTTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTCCAAATGGATGTT 1257
 DB |||||
 QY 721 CTAGTGCATGTTTACGTGGGATCTCCCGCTCGCCACCATCGCTATCGCTACATCATG 780
 DB |||||
 QY 1258 CTAGTGCATGTTTACGTGGGATCTCCCGCTCGCCACCATCGCTATCGCTACATCATG 1317
 DB |||||
 QY 781 AAGAGCATGCATGCTTTTACATGAGCTTACAGATTTGTGAAAGAAAAAGACCTACT 840
 DB |||||
 QY 1318 AAGAGCATGCATGCTTTTACATGAGCTTACAGATTTGTGAAAGAAAAAGACCTACT 1377
 DB |||||
 QY 841 ATATCTCCAAATCTCAAAATTTCTGGGCCAACTCTCTGGACTATGAGAAAGAAATTAAGAC 900
 DB |||||
 QY 1378 ATATCTCCAAATCTCAAAATTTCTGGGCCAACTCTCTGGACTATGAGAAAGAAATTAAGAC 1437
 DB |||||
 QY 901 CAGACT 906
 DB 1438 CAGACT 1443

PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3419; 11750bp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
Query Match 99.64; Score 902.8; DB 5; Length 5145;
Best Local Similarity 99.84; Pred. No. 8e-278; Indels 0; Gaps 0;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGTGAATTGGAAGTCAATTTGTTACTGAGAGTGTGCTGCTGCGAA 60
DB 589 ATGGCCCATGAGTGAATTGGAAGTCAATTTGTTACTGAGAGTGTGCTGCTGCGAA 648
QY 61 AGTGAACGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 649 AGTGAACGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
QY 121 CACATTTTGGAGCCATTATATCACTGCTCCAGGCTTGAAGGGAAGTGGCAACAG 180
DB 709 CACATTTTGGAGCCATTATATCACTGCTCCAGGCTTGAAGGGAAGTGGCAACAG 768
QY 181 GACAAAGTGTATATACAGAGCTCATCCAGCATTCAGCGAAGCATAGTGAATTGAT 240
DB 769 GACAAAGTGTATATACAGAGCTCATCCAGCATTCAGCGAAGCATAGTGAATTGAT 828
QY 241 TGCAGTCAGAGGTTGATTACGATCAAGGCTCCCAAGATGTTGCTCTCTCTCA 300
DB 829 TGCAGTCAGAGGTTGATTACGATCAAGGCTCCCAAGATGTTGCTCTCTCTCA 888
QY 301 GACTGTTTCTCACTGACTCTGGGTAACGTAAGAGGCTTCAACTGTTCACCTG 360
DB 889 GACTGTTTCTCACTGACTCTGGGTAACGTAAGAGGCTTCAACTGTTCACCTG 948
QY 361 CTTCGAGAGTGGTGTGCTGAGTCTCTGCTGTTTCCCTGGCCCTCTGTGAAGAAATCC 420
DB 949 CTTCGAGAGTGGTGTGCTGAGTCTCTGCTGTTTCCCTGGCCCTCTGTGAAGAAATCC 1008
QY 421 ACTCTAGTCCCTTACCTGATTTCTCAGCCCTTGGCTTACCTTGGCAACATGGGCCAAC 480
DB 1009 ACTCTAGTCCCTTACCTGATTTCTCAGCCCTTGGCTTACCTTGGCAACATGGGCCAAC 1068
QY 481 GGAATTTTCCCATCTTTATCTTGGCTGCAAGAGATGCTTCAACAGAGAGCTGATA 540
DB 1069 GGAATTTTCCCATCTTTATCTTGGCTGCAAGAGATGCTTCAACAGAGAGCTGATA 1128
QY 541 CAGCAGATGGATGTTGTTATGTTTAAATGCAAGCTATACTGTCCAAAGCTGACTTT 600
DB 1129 CAGCAGATGGATGTTGTTATGTTTAAATGCAAGCTATACTGTCCAAAGCTGACTTT 1188

QY 601 ATCCCGAGTCTCATTTCCGCGTGGCCGTGATGACAGCTTTTGTGAGAAATTTTG 660
DB 1189 ATCCCGAGTCTCATTTCCGCGTGGCCGTGATGACAGCTTTTGTGAGAAATTTTG 1248
QY 661 CCGTGGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCAATGATGTT 720
DB 1249 CCGTGGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCAATGATGTT 1308
QY 721 CTATGCTACTGTTTACTGGGATCTCCGCTCCGCCCAACCATCGCTATCGCTACATG 780
DB 1309 CTATGCTACTGTTTACTGGGATCTCCGCTCCGCCCAACCATCGCTATCGCTACATG 1368
QY 781 AAGAGATGACATGCTTTAGATGAGAGCTTACAGATTTGTGAAGAAAAAGACCTACT 840
DB 1369 AAGAGATGACATGCTTTAGATGAGAGCTTACAGATTTGTGAAGAAAAAGACCTACT 1428
QY 841 ATATCTCAAACTTCAATTTTCTGGCCCAACTCTGACTATGAGAAAGATTAAAGAC 900
DB 1429 ATATCTCAAACTTCAATTTTCTGGCCCAACTCTGACTATGAGAAAGATTAAAGAC 1488
QY 901 CAGACT 906
DB 1489 CAGACT 1494
RESULT 12
ABV21080
ID ABV21080 standard; cDNA; 5145 BP.
XX
XX ABV21080;
AC
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 21071.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; 88.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001MO-US005171.
XX
PR 17-FEB-2000; 2000US-0183119P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3481; 11750bp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 99.6%; Score 902.8; DB 5; Length 5145;
Best Local Similarity 99.8%; Pred. No. 8e-278; Indels 0; Gaps 0;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGCCCATGAGATGATGGAATCTCAATTGTTACTGAGAGTGTGCTGCTGGAA 60
DB 589 ATGGCCCATGAGATGATGGAATCTCAATTGTTACTGAGAGTGTGCTGCTGGAA 648
QY 61 AGTGAACGGAAAAAGTCTGCTAATTGATAGCCGCCATTGTGGAAATCAATACATCC 120
DB 649 AGTGAACGGAAAAAGTCTGCTAATTGATAGCCGCCATTGTGGAAATCAATACATCC 708
QY 121 CACATTTTGAAGCCATTATATATCACTGCTCAAGCTTATGAAGCGAAGTTGCAACG 180
DB 709 CACATTTTGAAGCCATTATATATCACTGCTCAAGCTTATGAAGCGAAGTTGCAACG 768
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGATTACGCGAAACATTAAGTTGACATTGAT 240
DB 769 GACAAAGTGTAAATTAACAGAGCTCATCCACATTACGCGAAACATTAAGTTGACATTGAT 828
QY 241 TGCAGTCAGAAAGTTGAGTTTACGATCAAGAGTCCCAAGATGTTGCCCTCTCTCTTA 300
DB 829 TGCAGTCAGAAAGTTGAGTTTACGATCAAGAGTCCCAAGATGTTGCCCTCTCTCTTA 888
QY 301 GACTGTTTCTCACTGACTTCTGGGTTAACTGGAGAAGAGTTCAACTCTGTTCACTG 360
DB 889 GACTGTTTCTCACTGACTTCTGGGTTAACTGGAGAAGAGTTCAACTCTGTTCACTG 948
QY 361 CTTGCAAGTGGGTTTGTGAGTCTCTGTTGTTCCCTGGGCTCGGAAAGAAATCC 420
DB 949 CTTGCAAGTGGGTTTGTGAGTCTCTGTTGTTCCCTGGGCTCGGAAAGAAATCC 1008
QY 421 ACTCTGTCCCTTACCTGCAATTTCTCAGCTTGTACTGTTGCCAAGATTGGGCCAAC 480
DB 1009 ACTCTGTCCCTTACCTGCAATTTCTCAGCTTGTACTGTTGCCAAGATTGGGCCAAC 1068
QY 481 CGAATTTCTTCCCATCTTATCTTGAGTCTCCAGAGATGTCCTCAACAGAGCTGATA 540
DB 1069 CGAATTTCTTCCCATCTTATCTTGAGTCTCCAGAGATGTCCTCAACAGAGCTGATA 1128
QY 541 CAGCGAATGGAATGTTGTTAATGTTAAATGCGAGCTTACCTGTCCAAAGCTGACTTT 600
DB 1129 CAGCGAATGGAATGTTGTTAATGTTAAATGCGAGCTTACCTGTCCAAAGCTGACTTT 1188
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 1189 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1248
QY 661 CCGTGGTGGACAAATCGTATGATTTCTTGAAGAAAGAAAGCTTCAATGATGTT 720
DB 1249 CCGTGGTGGACAAATCGTATGATTTCTTGAAGAAAGAAAGCTTCAATGATGTT 1308
QY 721 CTAGTGCACTGTTTATGCTGGAGTCTCCGCGCTCCGCGCAATGCTTATGCTTCAATC 780
DB 1309 CTAGTGCACTGTTTATGCTGGAGTCTCCGCGCTCCGCGCAATGCTTATGCTTCAATC 1368
QY 781 AAGAGATGAGCATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT 840
DB 1369 AAGAGATGAGCATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT 1428
QY 841 ATATCTCCAACTTCAATTTTCTGGGCGCACTCTGCACTATGAGAAAGATTGAAGAC 900
DB 1429 ATATCTCCAACTTCAATTTTCTGGGCGCACTCTGCACTATGAGAAAGATTGAAGAC 1488
QY 901 CAGACT 906
```

DB 1489 CAGACT 1494

RESULT 13

ABV26680
ID ABV26680 standard; cDNA, 5145 BP.

XX
AC ABV26680;

XX
DT 16-SEP-2002 (first entry)

XX
DE Human prostate expression marker cDNA 26671.

XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX
OS Homo sapiens.

XX
FN WO200160860-A2.

XX
PD 23-AUG-2001.

XX
PF 20-FEB-2001; 2001NO-US005171.

XX
PR 17-FEB-2000; 2000US-0183319P.

XX
PR 16-MAR-2000; 2000US-0189862P.

XX
PR 25-MAY-2000; 2000US-0207454P.

XX
PR 09-JUN-2000; 2000US-0211314P.

XX
PR 18-JUL-2000; 2000US-0219007P.

XX
PR 13-DEC-2000; 2000US-0255281P.

XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
PI Schlegel R, Endege WO, Monahan JE;

XX
DR WPI; 2001-662795/76.

XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX
PS Claim 1; Page 5388-5389; 11750P; English.

XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 99.6%; Score 902.8; DB 5; Length 5145;
Best Local Similarity 99.8%; Pred. No. 8e-278; Indels 0; Gaps 0;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGCCCATGAGATGATGGAATCTCAATTGTTACTGAGAGTGTGCTGCTGGAA 60
DB 589 ATGGCCCATGAGATGATGGAATCTCAATTGTTACTGAGAGTGTGCTGCTGGAA 648
QY 61 AGTGAACGGAAAAAGTCTGCTAATTGATAGCCGCCATTGTGGAAATCAATACATCC 120
DB 649 AGTGAACGGAAAAAGTCTGCTAATTGATAGCCGCCATTGTGGAAATCAATACATCC 708
QY 121 CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATGAGCGAAGTTGCAACG 180
```

DB 709 CACATTTGGAGCATTAATATCACTGCTCCAGCTTATGAAGGAGTTGCAACAG 768
QY 181 GACAAAGTGTATTTACAGAGCTATCCAGCATTTACAGGAAACATAAGTTGACATTGAT 240
DB 769 GACAAAGTGTATTTACAGAGCTATCCAGCATTTACAGGAAACATAAGTTGACATTGAT 828
QY 241 TGCAGTCAGAAAGTTGATTTAGATGATCAAAAGCTCCCAAGATGTTGGCTCTCTCTCA 300
DB 829 TGCAGTCAGAAAGTTGATTTAGATGATCAAAAGCTCCCAAGATGTTGGCTCTCTCTCA 888
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGACTTCACTCTGTTCACTG 360
DB 889 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGACTTCACTCTGTTCACTG 948
QY 361 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAAGAAATCC 420
DB 949 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAAGAAATCC 1008
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTGCAACATTTGGGCCAAC 480
DB 1009 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTGCAACATTTGGGCCAAC 1068
QY 481 CGAATTTCTTCCATCTTTATCTTGTGCTGACAGAGATGCTTCAACAGAGCTGATA 540
DB 1069 CGAATTTCTTCCATCTTTATCTTGTGCTGACAGAGATGCTTCAACAGAGCTGATA 1128
QY 541 CAGCAGAAATGGATGGTTATGTTTAAATGCCGCTATACCTGTCCAAAGCTGACTTT 600
DB 1129 CAGCAGAAATGGATGGTTATGTTTAAATGCCGCTATACCTGTCCAAAGCTGACTTT 1188
QY 601 ATCCCGAGTCTCATTTCTGCGGTCCTGTGAATGACAGCTTTGTGAGAAATTTTG 660
DB 1189 ATCCCGAGTCTCATTTCTGCGGTCCTGTGAATGACAGCTTTGTGAGAAATTTTG 1248
QY 661 CCGTGTGTGACAAATCAGTAGATTTCTATGAGAAAGCAAAAGCTTCCAAATGATGTT 720
DB 1249 CCGTGTGTGACAAATCAGTAGATTTCTATGAGAAAGCAAAAGCTTCCAAATGATGTT 1308
QY 721 CTATGTCATCTTTAGTGGGATCTCCGCTCCGACCAATCGCTATCGCTTACATGATG 780
DB 1309 CTATGTCATCTTTAGTGGGATCTCCGCTCCGACCAATCGCTATCGCTTACATGATG 1368
QY 781 AAGAGGATGGACATGCTTATGATGAGCTTACAGATTTGGAAGAAAAAGCTTACT 840
DB 1369 AAGAGGATGGACATGCTTATGATGAGCTTACAGATTTGGAAGAAAAAGCTTACT 1428
QY 841 ATATCTCAAACTTCAATTTCTGGGCCAACTCCTGGAATATGAGAAAGATTAAAGAC 900
DB 1429 ATATCTCAAACTTCAATTTCTGGGCCAACTCCTGGAATATGAGAAAGATTAAAGAC 1488
QY 901 CAGACT 906
DB 1489 CAGACT 1494
RESULT 14
ABV20978
ID ABV20978 standard; cDNA; 5145 BP.
XX AC ABV20978;
XX DT 13-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 20969.
XX KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KM pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX DR Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1, Page 3451; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
Query Match 99.6%; Score 902.8; DB 5; Length 5145;
Best Local Similarity 99.8%; Pred. No. 8e-278; 2; Indels 0; Gaps 0;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
DB 589 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 648
QY 61 AGTGAACGGAAGAGTCTGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 649 AGTGAACGGAAGAGTCTGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 708
QY 121 CACATTTGGAACCTTAAATATCACTGCTCAAGCTTATGAGGGAAGTTGCAACAG 180
DB 709 CACATTTGGAACCTTAAATATCACTGCTCAAGCTTATGAGGGAAGTTGCAACAG 768
QY 181 GACAAAGTGTATTTACAGAGCTATCCAGCATTTACAGGAAACATAAGTTGACATTGAT 240
DB 769 GACAAAGTGTATTTACAGAGCTATCCAGCATTTACAGGAAACATAAGTTGACATTGAT 828
QY 241 TGCAGTCAGAAAGTTGATTTAGATGATCAAAAGCTCCCAAGATGTTGGCTCTCTCTCA 300
DB 829 TGCAGTCAGAAAGTTGATTTAGATGATCAAAAGCTCCCAAGATGTTGGCTCTCTCTCA 888
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGACTTCACTCTGTTCACTG 360
DB 889 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGACTTCACTCTGTTCACTG 948
QY 361 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAAGAAATCC 420
DB 949 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAAGAAATCC 1008
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTGCAACATTTGGGCCAAC 480
DB 1009 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTGCAACATTTGGGCCAAC 1068

QY 481 CGAATCTCCCAATCTTATCTTGGCTGCGACGAGATGCTCAACAAGAGCTGATA 540
DB 1069 CGAATCTCTCCCAATCTTATCTTGGCTGCGACGAGATGCTCAACAAGAGCTGATA 1128
QY 541 CAGCAGAAATGGGATTTGTTATGTTAAATGCGAGTATCTGCTCAAGCTGACTT 600
DB 1129 CAGCAGAAATGGGATTTGTTATGTTAAATGCGAGTATCTGCTCAAGCTGACTT 1188
QY 601 ATCCCGAGTCTCAATCTCTGCGTGGCTGCTGATGACAGCTTTTGAGAAAATTTTG 660
DB 1189 ATCCCGAGTCTCAATCTCTGCGTGGCTGCTGATGACAGCTTTTGAGAAAATTTTG 1248
QY 661 CCGTGTGGGCAAAATAGATGATTTCAATGAGAAAGCAAAAGCTCCCAATGAGTGT 720
DB 1249 CCGTGTGGGCAAAATAGATGATTTCAATGAGAAAGCAAAAGCTCCCAATGAGTGT 1308
QY 721 CTAGTGACCTGTTAGCTGGGATCTCCGCTCGCCACCATGCTATGCGCTACATCATG 1368
DB 1309 CTAGTGACCTGTTAGCTGGGATCTCCGCTCGCCACCATGCTATGCGCTACATCATG 1368
QY 781 AAGAGATGACATGCTTTAATGATGAGCTTACAGATTTGAGAAAAGAAAGCTTACT 840
DB 1369 AAGAGATGACATGCTTTAATGATGAGCTTACAGATTTGAGAAAAGAAAGCTTACT 1428
QY 841 ATATCTCAAACTTCAATTTCTGCGGCACTCCCTGACATAGAGAAAGATTTAGAAC 900
DB 1429 ATATCTCAAACTTCAATTTCTGCGGCACTCCCTGACATAGAGAAAGATTTAGAAC 1488
QY 901 CAGACT 906
DB 1489 CAGACT 1494

RESULT 15
ABV21092
ID ABV21092 standard; cDNA: 5145 BP.

AC ABV21092;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 21083.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; se.
XX Homo sapiens.
XX OS
XX PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-021314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JB;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3485; 11750pp; English.
XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient; (e)
CC selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC / (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 99.6%; Score 902.8; DB 5; Length 5145;
Best Local Similarity 99.8%; Pred. No. 8e-278;

Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAAGTCAATGTTGTTACTGAGAGGTGAGCTCTGCTGAA 60
DB 589 ATGGCCCATGAGATGATGGAAGTCAATGTTGTTACTGAGAGGTGAGCTCTGCTGAA 648
QY 61 AGTGAACGGAAGAAAGTGTGCTGCTAATGATGATGACCGGCAATTTGTGAAATACATATCC 120
DB 649 AGTGAACGGAAGAAAGTGTGCTAATGATGATGATGACCGGCAATTTGTGAAATACATATCC 708
QY 121 CACATTTGGAAAGCCATTATATCAATGCTCCAGCTTATGAGAGCAAGGTGCAACAG 180
DB 709 CACATTTGGAAAGCCATTATATCAATGCTCCAGCTTATGAGAGCAAGGTGCAACAG 768
QY 181 GACAAAGTGTATTAACAGAGCTATCCAGATTCACGAAACATTAAGTTGACATTGAT 240
DB 769 GACAAAGTGTATTAACAGAGCTATCCAGATTCACGAAACATTAAGTTGACATTGAT 828
QY 241 TCGAGTCGGAAGCTGTGATTTAGATCAAGCTCCCAAGATGTCCTCTCTCTTCA 300
DB 829 TCGAGTCGGAAGCTGTGATTTAGATCAAGCTCCCAAGATGTCCTCTCTCTTCA 888
QY 301 GACTGTTTCTCACTGATCTTGGGTAACCTGAGAGCTTCAACTCTGTTCACTTG 360
DB 889 GACTGTTTCTCACTGATCTTGGGTAACCTGAGAGCTTCAACTCTGTTCACTTG 948
QY 361 CTTCGAGGTGGGTTTGTGAGTTCTCTGTTTCCCTGGGCTCTGTAAGAAATATCC 420
DB 949 CTTCGAGGTGGGTTTGTGAGTTCTCTGTTTCCCTGGGCTCTGTAAGAAATATCC 1008
QY 421 ACTTAATGCTTACCTGATTTCTCAGCTTGTCTTACCGTTGGCAATTTGGGCAACC 480
DB 1009 ACTTAATGCTTACCTGATTTCTCAGCTTGTCTTACCGTTGGCAATTTGGGCAACC 1068
QY 481 CGAATCTCCCAATCTTATCTTGGCTGCGACGAGATGCTCAACAAGAGCTGATA 540
DB 1069 CGAATCTCTCCCAATCTTATCTTGGCTGCGACGAGATGCTCAACAAGAGCTGATA 1128
QY 541 CAGCAGAAATGGGATTTGTTATGTTAAATGCGAGTATCTGCTCAAGCTGACTT 600
DB 1129 CAGCAGAAATGGGATTTGTTATGTTAAATGCGAGTATCTGCTCAAGCTGACTT 1188
QY 601 ATCCCGAGTCTCAATCTCTGCGTGGCTGCTGATGACAGCTTTTGAGAAAATTTTG 660
DB 1189 ATCCCGAGTCTCAATCTCTGCGTGGCTGCTGATGACAGCTTTTGAGAAAATTTTG 1248
QY 661 CCGTGTGGGCAAAATAGATGATTTCAATGAGAAAGCAAAAGCTCCCAATGAGTGT 720
DB 1249 CCGTGTGGGCAAAATAGATGATTTCAATGAGAAAGCAAAAGCTCCCAATGAGTGT 1308
QY 721 CTAGTGACCTGTTAGCTGGGATCTCCGCTCGCCACCATGCTATGCGCTACATCATG 780
DB 1309 CTAGTGACCTGTTAGCTGGGATCTCCGCTCGCCACCATGCTATGCGCTACATCATG 1368
QY 781 AAGAGATGACATGCTTTAATGATGAGCTTACAGATTTGAGAAAAGAAAGCTTACT 840

Db	1369	AAGAGATGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAAGACCTACT	1428
Qy	841	ATATCTCCAAACTTTCATTTTCTGGCCCACTCTGGACTATGAGAAGATTAGAAC	900
Db	1429	ATATCTCCAAACTTTCATTTTCTGGCCCACTCTGGACTATGAGAAGATTAGAAC	1488
Qy	901	CAGACT 906	
Db	1489	CAGACT 1494	

Search completed: June 21, 2004, 18:35:46
Job time : 528.548 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 18:08:06 ; Search time 540.914 Seconds
(without alignments)
7673.060 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443

Perfect score: 906
Sequence: 1 atggcccaatgagatgattg9.....agaagataagaccagact 906

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	902.8	99.6	1998	6	US-09-816-494-3
2	902.8	99.6	1998	16	US-10-377-072-27
3	902.8	99.6	2071	13	US-10-072-012-257
4	902.8	99.6	2732	13	US-10-168-506-2
5	902.8	99.6	3059	17	US-10-257-026-1
6	902.8	99.6	3496	9	US-09-964-277-1
7	902.8	99.6	3544	9	US-09-816-494-1
8	902.8	99.6	3544	16	US-10-377-072-25
9	902.8	99.6	3625	13	US-10-425-114-26234
10	902.8	99.6	3766	13	US-10-343-357-17
11	902.8	99.6	4790	17	US-10-648-593-115
12	901.2	99.5	2102	16	US-10-094-749-673
13	899.6	99.3	2966	13	US-10-296-115-520
14	817.8	90.3	2200	13	US-10-072-012-255

15	564.8	62.3	3332	9	US-09-964-277-20	Sequence 20, Appl
16	363.6	40.1	2453	14	US-10-005-858-1	Sequence 1, Appl
17	335.8	37.1	2476	13	US-10-220-120-28	Sequence 28, Appl
18	290.8	32.1	2039	13	US-10-072-012-265	Sequence 265, Appl
19	272.4	30.1	877	9	US-09-764-853-158	Sequence 158, Appl
20	223	24.6	277	17	US-10-637-855-29	Sequence 29, Appl
21	223	24.6	279	16	US-10-305-720-91	Sequence 91, Appl
22	217	24.0	246	17	US-10-257-026-3	Sequence 3, Appl
23	124.6	13.8	1337	9	US-09-964-899-46	Sequence 46, Appl
24	124.6	13.8	1449	12	US-10-052-482-152	Sequence 192, Appl
25	124.6	13.8	1635	17	US-10-648-593-27	Sequence 27, Appl
26	124.6	13.8	1830	15	US-10-346-356-1	Sequence 1, Appl
27	124.6	13.8	1909	12	US-10-052-482-191	Sequence 191, Appl
28	124.6	13.8	1909	13	US-10-058-482-129	Sequence 129, Appl
29	124.6	13.8	3286	10	US-09-971-329-9	Sequence 9, Appl
30	115	12.7	1452	12	US-10-052-482-189	Sequence 189, Appl
31	115	12.7	1977	12	US-10-052-482-188	Sequence 188, Appl
32	110.6	12.2	1356	15	US-10-184-832-6	Sequence 6, Appl
33	110.6	12.2	2741	15	US-10-184-832-4	Sequence 4, Appl
34	106.8	11.8	2303	9	US-09-880-107-3849	Sequence 3849, Ap
35	106.8	11.8	2303	10	US-09-922-146-3	Sequence 3, Appl
36	106.8	11.8	2303	13	US-10-342-887-603	Sequence 603, Appl
37	106.8	11.8	2303	13	US-10-172-118-603	Sequence 603, Appl
38	106.8	11.8	2338	13	US-10-363-616-7	Sequence 7, Appl
39	106.6	11.8	1143	15	US-10-184-832-3	Sequence 3, Appl
40	106.6	11.8	1574	13	US-10-296-115-440	Sequence 440, Appl
41	106.6	11.8	2650	15	US-10-184-832-1	Sequence 1, Appl
42	103.4	11.4	2104	12	US-10-152-318A-2041	Sequence 2041, Ap
43	103.4	11.4	2104	16	US-10-388-934-654	Sequence 654, Appl
44	103.4	11.4	2104	16	US-10-191-803-270	Sequence 270, Appl
45	99.2	10.9	462	10	US-09-918-995-31436	Sequence 31436, A

ALIGNMENTS

RESULT 1
US-09-816-494-3
Sequence 3, Application US/09816494
Patent No. US20020034807A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OR INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
TITLE OR INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,858
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1998
TYPE: DNA
ORGANISM: Homo sapiens
US-09-816-494-3

Query Match 99.6%; Score 902.8; DB 9; Length 1998;
Best Local Similarity 99.6%; Pred. No. 1,4e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGGCCCATGAGTGAATGTAAGTGAAGTGGTGGCTGCTGGAA	60
DB	1	ATGGCCCATGAGTGAATGTAAGTGAAGTGGTGGCTGCTGGAA	60
QY	61	AGTGAACGGAAGAGTGTCTGTAATGATAGCGGCATTTGTGATCAATACATCC	120
DB	61	AGTGAACGGAAGAGTGTCTGTAATGATAGCGGCATTTGTGATCAATACATCC	120
QY	121	CACATTTGGAAGCCATTAATCACTGCTCAAGCTTATGAAGCAAGTTGCAACAG	180
DB	121	CACATTTGGAAGCCATTAATCACTGCTCAAGCTTATGAAGCAAGTTGCAACAG	180


```
QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTGACATTGAT 240
Db 181 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTGACATTGAT 240
QY 241 TGCAGTCAGAAAGGTGTAGTATTAACAGATCAAGAGCTCCCAAGATGTGCTCTCTCTCA 300
Db 241 TGCAGTCAGAAAGGTGTAGTATTAACAGATCAAGAGCTCCCAAGATGTGCTCTCTCTCA 300
QY 301 GACTGTTTTCTCACTGATCTTCTGGGTAAACTGAGAAAGAGCTTCAACTCTGTCACTG 360
Db 301 GACTGTTTTCTCACTGATCTTCTGGGTAAACTGAGAAAGAGCTTCAACTCTGTCACTG 360
QY 361 CTGAGAGGTGGGTTTCTGAGATCTCTGTTGTTTCCCTGGGCTCTGTAAGAAATCC 420
Db 361 CTGAGAGGTGGGTTTCTGAGATCTCTGTTGTTTCCCTGGGCTCTGTAAGAAATCC 420
QY 421 ACTGTAGTCCCTACCTGATCTTCTGAGCTCCAGGCTTGTACTGTGCGCAACATTGGGCAAC 480
Db 421 ACTGTAGTCCCTACCTGATCTTCTGAGCTCCAGGCTTGTACTGTGCGCAACATTGGGCAAC 480
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTCCAGCGAGATGTCTCAAGAGAGCTGATA 540
Db 481 CGAATTTCTCCCAATCTTTATCTTGGCTCCAGCGAGATGTCTCAAGAGAGCTGATA 540
QY 541 CACAGAAATGGAGTGTATGTGTTAAATGCACTTACCTGTCCAAAGCCCTGACTT 600
Db 541 CACAGAAATGGAGTGTATGTGTTAAATGCACTTACCTGTCCAAAGCCCTGACTT 600
QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGATGACAGCTTTGTGAGAAATTTTG 660
Db 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGATGACAGCTTTGTGAGAAATTTTG 660
QY 661 CCGTGGTGGACAAATCAGATGATTTCACTGAGAAACAAAGCTTCCAAATGATGTT 720
Db 661 CCGTGGTGGACAAATCAGATGATTTCACTGAGAAACAAAGCTTCCAAATGATGTT 720
QY 721 CTAGTGCACGTGTTAGCTGGGATCTCCGCTCCGCCACATCCGCTATGCTCAATCANG 780
Db 721 CTAGTGCACGTGTTAGCTGGGATCTCCGCTCCGCCACATCCGCTATGCTCAATCANG 780
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAGCTTACT 840
Db 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAGCTTACT 840
QY 841 ATATCTCAAACTTCATTTTCTGGGCAACTCTCGAGCTATGAGAAAGAGTTAAAGAC 900
Db 841 ATATCTCAAACTTCATTTTCTGGGCAACTCTCGAGCTATGAGAAAGAGTTAAAGAC 900
QY 901 CAGACT 906
Db 901 CAGACT 906
```

RESULT 2

```
US-10-377-072-27
; Sequence 27, Application US/10377072
; Publication No. US2004009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; FILE REFERENCE: MD103-0180NM1M
; CURRENT APPLICATION NUMBER: US/10/377,072
; CURRENT FILING DATE: 2003-02-27
```

```

; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)
US-10-377-072-27
```

```
Query Match 99.6%; Score 902.8; DB 16; Length 1998;
Best Local Similarity 99.8%; Pred. No. 1.4e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ATGGGCCATGAGATGATTTGAACTCAATTTGTAAGTGAAGTGTGCTGCTGGGAA 60
Db 1 ATGGGCCATGAGATGATTTGAACTCAATTTGTAAGTGAAGTGTGCTGCTGGGAA 60
QY 61 AGTGAACGGAAGAAAGTCTCTAATTTGATAGCGCGCATTTTGGATATCAATATCC 120
Db 61 AGTGAACGGAAGAAAGTCTCTAATTTGATAGCGCGCATTTTGGATATCAATATCC 120
QY 121 CACATTTTGAAGCATTAATATCACTGCTCCAGCTTATGAGGAAAGTTGGAAG 180
Db 121 CACATTTTGAAGCATTAATATCACTGCTCCAGCTTATGAGGAAAGTTGGAAG 180
QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTGACATTGAT 240
Db 181 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTGACATTGAT 240
QY 241 TGCAGTCAGAAAGGTGTAGTATTAACAGATCAAGAGCTCCCAAGATGTGCTCTCTCTCA 300
Db 241 TGCAGTCAGAAAGGTGTAGTATTAACAGATCAAGAGCTCCCAAGATGTGCTCTCTCTCA 300
QY 301 GACTGTTTTCTCACTGATCTTCTGGGTAAACTGAGAAAGAGCTTCAACTCTGTCACTG 360
Db 301 GACTGTTTTCTCACTGATCTTCTGGGTAAACTGAGAAAGAGCTTCAACTCTGTCACTG 360
QY 361 CTGAGAGGTGGGTTTCTGAGATCTCTGTTGTTTCCCTGGGCTCTGTAAGAAATCC 420
Db 361 CTGAGAGGTGGGTTTCTGAGATCTCTGTTGTTTCCCTGGGCTCTGTAAGAAATCC 420
QY 421 ACTGTAGTCCCTACCTGATCTTCTGAGCTCCAGGCTTGTACTGTGCGCAACATTGGGCAAC 480
Db 421 ACTGTAGTCCCTACCTGATCTTCTGAGCTCCAGGCTTGTACTGTGCGCAACATTGGGCAAC 480
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTCCAGCGAGATGTCTCAAGAGAGCTGATA 540
Db 481 CGAATTTCTCCCAATCTTTATCTTGGCTCCAGCGAGATGTCTCAAGAGAGCTGATA 540
QY 541 CACAGAAATGGAGTGTATGTGTTAAATGCACTTACCTGTCCAAAGCCCTGACTT 600
```

Db	541	CAGCAGAAATGGGATTGGTTATGTTAAATGCCAGAAATCACTGTCCAAAGCTGACTTT	6000
OY	601	ATCCCCGAGTCTTCATTTCTCGCGTGTGCTGTGTAATGACAGCTTTTGTGAGAAATTTTG	6606
Db	601	ATCCCCGAGTCTTCATTTCTCGCGTGTGCTGTGTAATGACAGCTTTTGTGAGAAATTTTG	6606
OY	661	CCGTGGTTGGACAATACAGTAGATTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTGTT	7200
Db	661	CCGTGGTTGGACAATACAGTAGATTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTGTT	7200
OY	721	CTAGTGCACCTGTTTAGCTGGGATCTCCCGCTCCGCCACCAATCGCTATTCGCTTACATCATG	7800
Db	721	CTAGTGCACCTGTTTAGCTGGGATCTCCCGCTCCGCCACCAATCGCTTACATCATG	7800
OY	781	AAGAGGATGGAATGCTTTAGATGAAAGCTTCAGATTGTGAAAGAAAAAGACCTACT	8400
Db	781	AAGAGGATGGAATGCTTTAGATGAAAGCTTCAGATTGTGAAAGAAAAAGACCTACT	8400
OY	841	ATATCTCCAAACTTCATTTTCTGGGCCAACTCTGACTATGAGAGAAAGATTAAAGAAC	9000
Db	841	ATATCTCCAAACTTCATTTTCTGGGCCAACTCTGACTATGAGAGAAAGATTAAAGAAC	9000
OY	901	CAGACT 906	
Db	901	CAGACT 906	

RESULT 3

US-10-072-012-257
Sequence 257, Application US/10072012
Publication No. US2004003493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spyrek, Kimberly
APPLICANT: Zernusen, Bryan
APPLICANT: Patnurejan, Meera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Kalazyna
APPLICANT: Grosse, William M.
APPLICANT: Alsebrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767

	:	PRIOR FILING DATE:	2001-02-05	
	:	PRIOR APPLICATION NUMBER:	60/267,057	
	:	PRIOR FILING DATE:	2001-02-07	
	:	PRIOR APPLICATION NUMBER:	60/266,975	
	:	PRIOR FILING DATE:	2001-02-07	
	:	PRIOR APPLICATION NUMBER:	60/267,459	
	:	PRIOR FILING DATE:	2001-02-08	
	:	Remaining Prior Application data removed - See File Wrapper or PALM.		
	:	NUMBER OF SEQ ID NOS:	1391	
	:	SOFTWARE:	PatentIn Ver. 2.1	
	:	SEQ ID NO:	257	
	:	LENGTH:	2071	
	:	TYPE:	DNA	
	:	ORGANISM:	Homo sapiens	
	:	US-10-072-012-257		
		Query Match	99.6%; Score 902.8; DB 13; Length 2071;	
		Best Local Similarity	99.8%; Pred. No. 1.5e-282;	
		Matches	904; Conservative 0; Mismatches 2; Indels 0; Gaps 0	
OY		1	ATGGCCCATGAGTGAATGTGAACTCAATTGTACTGAGAGTTGGTGCTCTGCTGGAA	60
DB		61	ATGGCCCCATGAGATGATTGAACCAATTGTACTGAGAGTTGGTGCTCTGCTGGAA	120
OY		61	AGTGGAAACGGAAAAAGTCGTCTAATTGATPAGCGGCACATTGTGTAATCAATAATCACC	120
DB		121	AGTGGAAACGGAAAAAGTCGTCTAATTGATPAGCGGCACATTGTGTAATCAATAATCACC	180
OY		121	CACATTTTGGAAAGCCATTAATATCAACTGCTCCAAGCTTATGAAACGAAGTTGGCAACG	180
DB		181	CACATTTTGGAAAGCCATTAATATCAACTGCTCCAAGCTTATGAAACGAAGTTGGCAACG	240
OY		181	GACAAAGTGTTAATTACAGAGCTCACAGCAATTCAGCGAAACATAAGTTGACATTGAT	240
DB		241	GACAAAGTGTTAATTACAGAGCTCACAGCAATTCAGCGAAACATAAGTTGACATTGAT	300
OY		241	TGCAAGTCAGAAGTTGTGTTTACGATCAAAGCTCCCAAGATGTTGCCCTCTCTCTTCA	300
DB		301	TGCAAGTCAGAAGTTGTGTTTACGATCAAAGCTCCCAAGATGTTGCCCTCTCTCTTCA	360
OY		301	GACTGTTTCTCACTGTACTCTCGGGTAACTGGAAAGAGCTTCAACTGTGTCACTCG	360
DB		361	GACTGTTTCTCACTGTACTCTCGGGTAACTGGAAAGAGCTTCAACTGTGTGTCACTCG	420
OY		361	CTTGCAGGTGGGTTTGTCTGAGTTCTCTCGTTGTTTCCCTGGCCTCTGTGAAGAAAATCC	420
DB		421	CTTGCAGGTGGGTTTGTCTGAGTTCTCTCGTTGTTTCCCTGGCCTCTGTGAAGAAAATCC	480
OY		421	ACTCTAGTCTCTACCTGATTTCTCAGCCTTGCTTAACCTGTGTCACAATTGGGCAACC	480
DB		481	ACTCTAGTCTCTACCTGATTTCTCAGCCTTGCTTAACCTGTGTCACAATTGGGCAACC	540
OY		481	GGAATTCCTCCCAATCTTTATCTTGAGCTGCGACGAGATGTCCTCAACAAAGAGCTGATA	540
DB		541	GGAATTCCTCCCAATCTTTATCTTGAGCTGCGACGAGATGTCCTCAACAAAGAGCTGATA	600
OY		541	CAGCAGATGAGGATTTGTTATGTGTTAAATGCCAGCTATACCTGTGTCAAAAGCTGACTT	600
DB		601	CAGCAGATGAGGATTTGTTATGTGTTAAATGCCAGCTATACCTGTGTGTCAAAAGCTGACTT	660
OY		601	ATCCCCGAGTCTCATTTCTCTGCGGTGTCCTGTAATGACAGCTTTTGTGAAGAAAATTTTG	660
DB		661	ATCCCCGAGTCTCATTTCTCTGCGGTGTCCTGTAATGACAGCTTTTGTGAAGAAAATTTTG	720
OY		661	CCGTGTTTGGCAAAATCAGTAGATTCTCATGAGAAAGCAAAAGCTCCAAATGAGATGTT	720
DB		721	CCGTGTTTGGCAAAATCAGTAGATTCTCATGAGAAAGCAAAAGCTCCAAATGAGATGTT	780
OY		721	CTAGTGCACTGTTTAGCTGGGATCTCCGCTCCGCAACCATGCTATGCGCTTACATCATG	780
DB		781	CTAGTGCACTGTTTAGCTGGGATCTCCGCTCCGCAACCATGCTATGCGCTTACATCATG	840
OY		781	AAGAGGATGACATGCTCTTTTAGATGAGACCTTACAGATTTGTGAAGAAAAGAACCTACT	840

Db 307 GACAAAGTGTAAATACAGAGCTCATCCAGCAATTCAGCAAAATAGTTGACATTCAT 366
Qy 241 TGCAGTCAGAAAGTTGATTTAGATTAAGATCAAGCTCCCAAGATTTGGCTCTCTCTTCA 300
Db 367 TGCAGTCAGAAAGTTGATTTAGATTAAGATCAAGCTCCCAAGATTTGGCTCTCTCTTCA 426
Qy 301 GACGTTTCTCACTGACTTCTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACTG 360
Db 427 GACTGTTTCTCACTGACTTCTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACTG 486
Qy 361 CTTCAGAGTGGTGTGCTGAGATTCCTGTTGTTCCCTGGCTCTGTTGAGAGAAATCC 420
Db 487 CTTCAGAGTGGTGTGCTGAGATTCCTGTTGTTCCCTGGCTCTGTTGAGAGAAATCC 546
Qy 421 ACTCTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACTGTTGCAACATTTGGCCCAAC 480
Db 547 ACTCTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACTGTTGCAACATTTGGCCCAAC 606
Qy 481 CGAATCTTCCCAATCTTTATCTTGGCTGCGAGAGATGCTTCAACAGAGCTGATA 540
Db 607 CGAATCTTCCCAATCTTTATCTTGGCTGCGAGAGATGCTTCAACAGAGCTGATA 666
Qy 541 CACAGAAATGGATGGTATGTTAAATGCAAGCTATACTGTCCAAAGCTGACTTT 600
Db 667 CACAGAAATGGATGGTATGTTAAATGCAAGCTATACTGTCCAAAGCTGACTTT 726
Qy 601 ATCCCGAGTCTCATTTCTGCTGTGCTGTGATGACAGCTTTGTGAGAAATTTTG 660
Db 727 ATCCCGAGTCTCATTTCTGCTGTGCTGTGATGACAGCTTTGTGAGAAATTTTG 786
Qy 661 CGGTGGTGGCAAAATAGATGATTTGATGAGAAAGCAAAAGCTTCAATGATGTT 720
Db 787 CGGTGGTGGCAAAATAGATGATTTGATGAGAAAGCAAAAGCTTCAATGATGTT 846
Qy 721 CTAGTCACTGTTTACCTGGATCTCCCGCTCGCCCACTATGCTTACATCATG 780
Db 847 CTAGTCACTGTTTACCTGGATCTCCCGCTCGCCCACTATGCTTACATCATG 906
Qy 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGCTTACT 840
Db 907 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGCTTACT 966
Qy 841 ATATCTCCAAACTTCAATTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
Db 967 ATATCTCCAAACTTCAATTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAGAAC 1026
Qy 901 CAGACT 906
Db 1027 CAGACT 1032

RESULT 6

US-09-964-277-1
; Sequence 1, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-1

Query Match 99.6%; Score 902.8; DB 9; Length 3496;
Best Local Similarity 99.8%; Pred. No. 26-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATGGAACTCAAAATTGTTACTGAGAGTTGGTGGCTCTGTCGAA 60
Db 562 ATGGCCCATGAGATGATGGAACTCAAAATTGTTACTGAGAGTTGGTGGCTCTGTCGAA 621
Qy 61 AGTGAACGGAAGAAAGTGTCTGATTAATGATAGCGGCCCAATTTGTGAAATACAAATAC 120
Db 622 AGTGAACGGAAGAAAGTGTCTGATTAATGATAGCGGCCCAATTTGTGAAATACAAATAC 681
Qy 121 CACATTTGGAAAGCCATTAATATCACTGCTCCAGCTTATGAAAGGAGTTGCAACAG 180
Db 682 CACATTTGGAAAGCCATTAATATCACTGCTCCAGCTTATGAAAGGAGTTGCAACAG 741
Qy 181 GACAAAGTGTAAATACAGAGCTCATCCAGCTTACGCGAAACATTAAGTTGACATTCAT 240
Db 742 GACAAAGTGTAAATACAGAGCTCATCCAGCTTACGCGAAACATTAAGTTGACATTCAT 801
Qy 241 TGCAGTCAGAAAGTTGATTTAGATTAAGATCAAGCTCCCAAGATTTGCTCTCTTCA 300
Db 802 TGCAGTCAGAAAGTTGATTTAGATTAAGATCAAGCTCCCAAGATTTGCTCTCTTCA 861
Qy 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAGAGACTTCAACTCTGTTCACTG 360
Db 862 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAGAGACTTCAACTCTGTTCACTG 921
Qy 361 CTTCAGAGTGGTGTGCTGAGATTCCTGTTGTTCCCTGGCTCTGTTGAGAGAAATCC 420
Db 922 CTTCAGAGTGGTGTGCTGAGATTCCTGTTGTTCCCTGGCTCTGTTGAGAGAAATCC 981
Qy 421 ACTCTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACTGTTGCAACATTTGGCCCAAC 480
Db 982 ACTCTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACTGTTGCAACATTTGGCCCAAC 1041
Qy 481 CGAATCTTCCCAATCTTTATCTTGGCTGCGAGAGATGCTTCAACAGAGCTGATA 540
Db 1042 CGAATCTTCCCAATCTTTATCTTGGCTGCGAGAGATGCTTCAACAGAGCTGATA 1101
Qy 541 CACAGAAATGGATGGTATGTTAAATGCAAGCTATACCTGTCCAAAGCTGACTTT 600
Db 1102 CACAGAAATGGATGGTATGTTAAATGCAAGCTATACCTGTCCAAAGCTGACTTT 1161
Qy 601 ATCCCGAGTCTCATTTCTGCTGTGCTGTGATGACAGCTTTGTGAGAAATTTTG 660
Db 1162 ATCCCGAGTCTCATTTCTGCTGTGCTGTGATGACAGCTTTGTGAGAAATTTTG 1221
Qy 661 CGGTGGTGGCAAAATAGATGATTTGATGAGAAAGCAAAAGCTTCAATGATGTT 720
Db 1222 CGGTGGTGGCAAAATAGATGATTTGATGAGAAAGCAAAAGCTTCAATGATGTT 1281
Qy 721 CTAGTCACTGTTTACCTGGATCTCCCGCTCGCCCACTATGCTTACATCATG 780
Db 1282 CTAGTCACTGTTTACCTGGATCTCCCGCTCGCCCACTATGCTTACATCATG 1341
Qy 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGCTTACT 840
Db 1342 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGCTTACT 1401
Qy 841 ATATCTCCAAACTTCAATTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAGAAC 900
Db 1402 ATATCTCCAAACTTCAATTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAGAAC 1461
Qy 901 CAGACT 906
Db 1462 CAGACT 1467

RESULT 7

US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY

```

; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589)...(2583)
US-09-816-494-1
```

```

Query Match          99.6%; Score 902.8; DB 9; Length 3544;
Best Local Similarity 99.6%; Pred. No. 2.1e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 ATGGCCCATGAGATGATGGAATCTCAATTTGTTACGAGAGTGGTGGCTCTGCTGAA 60
DB 589 ATGCCCCATGAGATGATGGAATCTCAATTTGTTACGAGAGTGGTGGCTCTGCTGAA 60
QY 61 AGTGAACGGAAAAAGTGTCTAATTGATAGCCGCCATTGTGGAATCAATATACATCC 120
DB 649 AGTGAACGGAAAAAGTGTCTAATTGATAGCCGCCATTGTGGAATCAATATACATCC 120
QY 121 CACATTTTGAAGCCATTAAATATCAACTGCTCAAGCTTATGAAAGCAAGTTGCAACG 180
DB 709 CACATTTTGAAGCCATTAAATATCAACTGCTCAAGCTTATGAAAGCAAGTTGCAACG 180
QY 181 GACAAAGTGTAAATTACAGAGCTCATCAGATTCAGCCAAACATAGGTTGACATGAT 240
DB 769 GACAAAGTGTAAATTACAGAGCTCATCAGATTCAGCCAAACATAGGTTGACATGAT 240
QY 241 TGGCTGGAAGTGTGATTTAGATCAAGTCCCAAGATGTTGCCCTCTCTCTTCA 300
DB 829 TGGCTGGAAGTGTGATTTAGATCAAGTCCCAAGATGTTGCCCTCTCTCTTCA 300
QY 301 GACTGTTTCTCACTGACTTCTGGGTAACAGGTAACAGGTAACAGGTAACAGGTAAC 360
DB 889 GACTGTTTCTCACTGACTTCTGGGTAACAGGTAACAGGTAACAGGTAACAGGTAAC 360
QY 361 CTGACAGTGGGTTTGTGAGTTCTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 949 CTGACAGTGGGTTTGTGAGTTCTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 ACTCTAGTCTCTACCTGATTTCTCAGCTTGTCTTACTGTTGCAACATTTGGCCCAAC 480
DB 1009 ACTCTAGTCTCTACCTGATTTCTCAGCTTGTCTTACTGTTGCAACATTTGGCCCAAC 480
QY 481 CGAATCTTCCCAATCTTATTTGGCTGCGCAGAGATGCTCTCAACAGAGGCTGAT 540
DB 1069 CGAATCTTCCCAATCTTATTTGGCTGCGCAGAGATGCTCTCAACAGAGGCTGAT 540
QY 541 CAGCAAGATGGATTTGTTATGTTAAAGCCAGTATCTCTGCAACAGAGGCTGAT 1128
DB 1129 CAGCAAGATGGATTTGTTATGTTAAAGCCAGTATCTCTGCAACAGAGGCTGAT 1128
QY 601 ATCCCGAGTCTCAATTTCTGCGTGGCTGCTGATATGACAGCTTTTGTAGAAATTTTG 660
DB 1189 ATCCCGAGTCTCAATTTCTGCGTGGCTGCTGATATGACAGCTTTTGTAGAAATTTTG 660
QY 661 CCGTGTGTTGAACAATCAGTATGATTTCAATTTGAGAAAGCAAAAGCTCCCAATGATGTT 720
DB 1249 CCGTGTGTTGAACAATCAGTATGATTTCAATTTGAGAAAGCAAAAGCTCCCAATGATGTT 720
QY 721 CTAGTGACCTTTAGCTGGAGTCTCCGCTCCGACCATGCTATGCTCTCAATCATGAT 780
DB 1309 CTAGTGACCTTTAGCTGGAGTCTCCGCTCCGACCATGCTATGCTCTCAATCATGAT 780
```

```

QY 781 AAGAGATGACATGCTTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 1369 AAGAGATGACATGCTTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCCACTCCGCACTATGAGAAAGATTAGAAC 900
DB 1429 ATATCTCCAACTTCAATTTTCTGGGCCCACTCCGCACTATGAGAAAGATTAGAAC 900
QY 901 CAGACT 906
DB 1489 CAGACT 1494
```

RESULT 8

```

US-10-377-072-25
; Sequence 25, Application US/10377072
; Publication No. US2004009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Currie, Kory A.J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Mayoung
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 38992, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; FILE REFERENCE: MP103-0180NM1M
; CURRENT APPLICATION NUMBER: US/10/377,072
; PRIOR FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589)...(2586)
```

```

US-10-377-072-25
```

```

Query Match          99.6%; Score 902.8; DB 16; Length 3544;
Best Local Similarity 99.6%; Pred. No. 2.1e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 ATGGCCCATGAGATGATGGAATCTCAATTTGTTACGAGAGTGGTGGCTCTGCTGAA 60
DB 589 ATGCCCCATGAGATGATGGAATCTCAATTTGTTACGAGAGTGGTGGCTCTGCTGAA 60
```

QY 61 AGTGAACGAAAAAGTGTGCTAATTGATAGCCGCCATTTGTGGAATCAATACATCC 120
Db 649 AGTGAACGAAAAAGTGTGCTAATTGATAGCCGCCATTTGTGGAATCAATACATCC 708
QY 121 CACATTTTGAAGCATTAAATATCACTGCTCCAAAGCTTAAGAGCAAGTTGCAACAG 180
Db 709 CACATTTTGAAGCATTAAATATCACTGCTCCAAAGCTTAAGAGCAAGTTGCAACAG 768
QY 181 GACAAAGTGTAAATTAAGAGCTCATCCAGCATTCACAGAAACATAAGTTGACATTGAT 240
Db 769 GACAAAGTGTAAATTAAGAGCTCATCCAGCATTCACAGAAACATAAGTTGACATTGAT 828
QY 241 TGCAGTCAGAAAGTGTAGTTTACATCAAGCTCCCAAGATGTGCTCTCTCTTCA 300
Db 829 TGCAGTCAGAAAGTGTAGTTTACATCAAGCTCCCAAGATGTGCTCTCTCTTCA 888
QY 301 GACTGTTTCTCACTGACTTCTGGGTAAACAGAGAAAGCTTCAACTCTGTTCACCTG 360
Db 889 GACTGTTTCTCACTGACTTCTGGGTAAACAGAGAAAGCTTCAACTCTGTTCACCTG 948
QY 361 CTTCGAGTGGGTTTGTGAGTTCTCTGCTGTTTCCCTGGCCCTGTGTGAAGAAATCC 420
Db 949 CTTCGAGTGGGTTTGTGAGTTCTCTGCTGTTTCCCTGGCCCTGTGTGAAGAAATCC 1008
QY 421 ACTCTAGTCCCTACCTGCTCATTTCTCAGCCTTGCTTACCTGTGCAACATTGGGCCAAC 480
Db 1009 ACTCTAGTCCCTACCTGCTCATTTCTCAGCCTTGCTTACCTGTGCAACATTGGGCCAAC 1068
QY 481 CGAATCTTCCCAATCTTTATCTTGGCTGCGCAGAGAGATGCTCAACAAAGAGCTGATA 540
Db 1069 CGAATCTTCCCAATCTTTATCTTGGCTGCGCAGAGAGATGCTCAACAAAGAGCTGATA 1128
QY 541 CACAGAAATGGGATGGTTATGTGTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 600
Db 1129 CACAGAAATGGGATGGTTATGTGTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 1188
QY 601 ATCCCGAGTCTCATTTCTCGCTGTGCTGTGAATGACAGCCTTTGTGAGAAATTTTG 660
Db 1189 ATCCCGAGTCTCATTTCTCGCTGTGCTGTGAATGACAGCCTTTGTGAGAAATTTTG 1248
QY 661 CCGTGGTTGGACAAATCAGTAGATTCTTATGAGAAAGCAAAAGCCTTCAATGATGTGT 720
Db 1249 CCGTGGTTGGACAAATCAGTAGATTCTTATGAGAAAGCAAAAGCCTTCAATGATGTGT 1308
QY 721 CTAGTGACCTGTTTAAAGTGGGATCTCCCGCTCCGCAACATCGCTATCGCTACATCATG 780
Db 1309 CTAGTGACCTGTTTAAAGTGGGATCTCCCGCTCCGCAACATCGCTATCGCTACATCATG 1368
QY 781 AAGAGATGAGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAGAGCCACT 840
Db 1369 AAGAGATGAGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAGAGCCACT 1428
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAATCCTGTGACTATGAGAAAGATTAAGAAC 900
Db 1429 ATATCTCCAAACTTCAATTTTCTGGGCCAATCCTGTGACTATGAGAAAGATTAAGAAC 1488
QY 901 CAGACT 906
Db 1489 CAGACT 1494

RESULT 9

US-10-425-114-26234
; Sequence 26234, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaka, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313) B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26234
; LENGTH: 3625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-028-H6_FLI
US-10-425-114-26234

Query Match 99.6%; Score 902.8; DB 13; Length 3625;
Best Local Similarity 99.8%; Pred. No. 2.1e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCATGAGATGATGGAATCTCAATTTGTTACTAGAGTGTGTGCTGTGCGAA 60
Db 692 ATGGCCATGAGATGATGGAATCTCAATTTGTTACTAGAGTGTGTGCTGTGCGAA 751
QY 61 AGTGAACGAAAAAGTGTGCTAATTGATAGCCGCCATTTGTGGAATCAATACATCC 120
Db 752 AGTGAACGAAAAAGTGTGCTAATTGATAGCCGCCATTTGTGGAATCAATACATCC 811
QY 121 CACATTTTGAAGCATTAAATATCACTGCTCCAAAGCTTAAGAGCAAGTTGCAACAG 180
Db 812 CACATTTTGAAGCATTAAATATCACTGCTCCAAAGCTTAAGAGCAAGTTGCAACAG 871
QY 181 GACAAAGTGTAAATTAAGAGCTCATCCAGCATTCACAGAAACATAAGTTGACATTGAT 240
Db 872 GACAAAGTGTAAATTAAGAGCTCATCCAGCATTCACAGAAACATAAGTTGACATTGAT 931
QY 241 TGCAGTCAGAAAGTGTAGTTTACATCAAGCTCCCAAGATGTGCTCTCTCTTCA 300
Db 932 TGCAGTCAGAAAGTGTAGTTTACATCAAGCTCCCAAGATGTGCTCTCTCTTCA 991
QY 301 GACTGTTTCTCACTGACTTCTGGGTAAACAGAGAAAGCTTCAACTCTGTTCACCTG 360
Db 992 GACTGTTTCTCACTGACTTCTGGGTAAACAGAGAAAGCTTCAACTCTGTTCACCTG 1051
QY 361 CTTCGAGTGGGTTTGTGAGTTCTCTGCTGTTTCCCTGGCCCTGTGTGAAGAAATCC 420
Db 1052 CTTCGAGTGGGTTTGTGAGTTCTCTGCTGTTTCCCTGGCCCTGTGTGAAGAAATCC 1111
QY 421 ACTTAGTCCCTACCTGCTCATTTCTCAGCCTTGCTTACCTGTGCAACATTGGGCCAAC 480
Db 1112 ACTTAGTCCCTACCTGCTCATTTCTCAGCCTTGCTTACCTGTGCAACATTGGGCCAAC 1171
QY 481 CGAATCTTCCCAATCTTTATCTTGGCTGCGCAGAGAGATGCTCAACAAAGAGCTGATA 540
Db 1172 CGAATCTTCCCAATCTTTATCTTGGCTGCGCAGAGAGATGCTCAACAAAGAGCTGATA 1231
QY 541 CACAGAAATGGGATGGTTATGTGTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 600
Db 1232 CACAGAAATGGGATGGTTATGTGTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 1291
QY 601 ATCCCGAGTCTCATTTCTCGCTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 660
Db 1292 ATCCCGAGTCTCATTTCTCGCTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 1351
QY 661 CCGTGGTTGGACAAATCAGTAGATTCTTATGAGAAAGCAAAAGCCTTCAATGATGTGT 720
Db 1352 CCGTGGTTGGACAAATCAGTAGATTCTTATGAGAAAGCAAAAGCCTTCAATGATGTGT 1411
QY 721 CTAGTGACCTGTTTAAAGTGGGATCTCCCGCTCCGCAACATCGCTATCGCTACATCATG 780
Db 1412 CTAGTGACCTGTTTAAAGTGGGATCTCCCGCTCCGCAACATCGCTATCGCTACATCATG 1471
QY 781 AAGAGATGAGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAGAGCCACT 840
Db 1472 AAGAGATGAGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAGAGCCACT 1531

QY 841 ATATCTCCAACTTCAATTTCTGGGCGCACTCTGCGATCTATGAGAGAGATTAAAGAC 900
DB 1532 ATATCTCCAACTTCAATTTCTGGGCGCACTCTGCGATCTATGAGAGAGATTAAAGAC 1591
QY 901 CAGACT 906
DB 1592 CAGACT 1597

RESULT 10
US-10-343-357-17
; Sequence 17, Application US/10343357
; Publication No. US20040058341A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y.Tom
; APPLICANT: ELIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
; APPLICANT: LEE, Ernestine A.; HARALIA, April J.A.
; APPLICANT: GRIFPIN, Jennifer A.; BAUGHEN, Catherine M.
; APPLICANT: YUE, Henry; WARREN, Bridget A.
; APPLICANT: NEUREN, Daniel B.; CHAWLA, Narinder K.
; APPLICANT: KEARNEY, Liam
; TITLE OF INVENTION: PROTEIN PHOSPHATASES
; FILE REFERENCE: PI-0173 PCT
; CURRENT APPLICATION NUMBER: US/10/343,357
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: PCT/US01/23716
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,679
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/223,272
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/224,309
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 60/226,728
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/229,254
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: US 60/231,366
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 3766
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CBI
US-10-343-357-17

Query Match 99.6%; Score 902.8; DB 13; Length 3766;
Best Local Similarity 99.8%; Pred. No. 2.1e-282;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAATCAATTTGTAATGAGAGGTTGGCTGCTGGAA 60
DB 538 ATGGCCCATGAGATGATGGAATCAATTTGTAATGAGAGGTTGGCTGCTGGAA 597
QY 61 AGTGAAGGAAAAGCTCTCTAATGATGAGCGGCAATTTGTAATGATGATGATGAT 120
DB 598 AGTGAAGGAAAAGCTCTCTAATGATGAGCGGCAATTTGTAATGATGATGATGAT 657
QY 121 CACATTTTGAAGCATTATATCACTGCTCAAGCTTATGAGCGAAGGTTGCAAC 180
DB 658 CACATTTTGAAGCATTATATCACTGCTCAAGCTTATGAGCGAAGGTTGCAAC 717
QY 181 GACAAAGTGTATATTCAGAGCTCATCCAGCAATTCAGCAAAATGAGGTTGATGAT 240
DB 718 GACAAAGTGTATATTCAGAGCTCATCCAGCAATTCAGCAAAATGAGGTTGATGAT 777

QY 241 TGCAGTCAAGAGGTTGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 778 TGCAGTCAAGAGGTTGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 837
QY 301 GACTGTTTCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 838 GACTGTTTCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 897
QY 361 CTGAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 898 CTGAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 957
QY 421 ACTGATGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 958 ACTGATGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1017
QY 481 CGAATTTCTCCCAATTTTATCTGATGATGATGATGATGATGATGATGATGATGAT 540
DB 1018 CGAATTTCTCCCAATTTTATCTGATGATGATGATGATGATGATGATGATGATGAT 1077
QY 541 CAGAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 1078 CAGAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1137
QY 601 ATCCCGAGTCTCATTTCTGATGATGATGATGATGATGATGATGATGATGAT 660
DB 1138 ATCCCGAGTCTCATTTCTGATGATGATGATGATGATGATGATGATGATGAT 1197
QY 661 CGGTGTTGCAAAATCAGTATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 1198 CGGTGTTGCAAAATCAGTATGATGATGATGATGATGATGATGATGATGATGAT 1257
QY 721 CTATGTCATGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 1258 CTATGTCATGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
QY 781 AAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 1318 AAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1377
QY 841 ATATCTCCAACTTCAATTTCTGGGCGCACTCTGCGATCTATGAGAGAGATTAAAGAC 900
DB 1378 ATATCTCCAACTTCAATTTCTGGGCGCACTCTGCGATCTATGAGAGAGATTAAAGAC 1437
QY 901 CAGACT 906
DB 1438 CAGACT 1443

RESULT 11
US-10-648-593-115
; Sequence 115, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 115
; LENGTH: 4790
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-648-593-115

Query Match 99.6%; Score 902.8; DB 17; Length 4790;

Query	Match	99.5%	Score	901.2	DB 16	Length	2102
Db	Best Local Similarity	99.7%	Pred. No.	4.9e-282			
Matches	903	Conservative	0	Mismatches	3	Indels	0
						Gaps	0
QY	1 ATGGCCCATGAGATGATTTGAACCTCAATTGTTACTGAGAGGTTGGCTCTGCTGGA	60					
Db	56 ATGGCCCATGAGATGATTTGAACCTCAATTGTTACTGAGAGGTTGGCTCTGCTGGA	115					
QY	61 AGTGAACGGAAGAGTCTGCTGAATTGATAGCCGCGCAATTTGTGGAATACATACATCC	120					
Db	116 AGTGAACGGAAGAGTCTGCTGAATTGATAGCCGCGCAATTTGTGGAATACATACATCC	175					
QY	121 CACATTTTGGAGGCATTAATATCAACGCTCCCAAGCTTATGAGCGAAGTTGACACAG	180					
Db	176 CACATTTTGGAGGCATTAATATCAACGCTCCCAAGCTTATGAGCGAAGTTGACACAG	235					
QY	181 GACAAAGTGAATTAACAGAGCTCATCAGCAATTCAGCGAAACATTAAGTTGACATTTGAT	240					
Db	236 GACAAAGTGAATTAACAGAGCTCATCAGCAATTCAGCGAAACATTAAGTTGACATTTGAT	295					
QY	241 TGCAGTCGAGAGGTTGATGATTACGATCAAGCTCCAAAGATGTTGCTCTCTTTCA	300					
Db	296 TGCAGTCGAGAGGTTGATGATTACGATCAAGCTCCAAAGATGTTGCTCTCTTTCA	355					
QY	301 GACTGTTTTCACATCTGTACTTCTGGGTAACTGAGAGAAAGCTTCAATCTCTGTTACCTG	360					
Db	356 GACTGTTTTCACATCTGTACTTCTGGGTAACTGAGAGAAAGCTTCAATCTCTGTTACCTG	415					
QY	361 CTTCGAGGTGGGTTTGTGAGGTTCTCTGCTGTTTCCCTGCGCTCTGTAAGGAATAATCC	420					
Db	416 CTTCGAGGTGGGTTTGTGAGGTTCTCTGCTGTTTCCCTGCGCTCTGTAAGGAATAATCC	475					
QY	421 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGCCTTACTGTTGCCCAATTTGGGCCAAC	480					
Db	476 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGCCTTACTGTTGCCCAATTTGGGCCAAC	535					
QY	481 CGAATTTCTTCCCATCTTTATCTTGGCTGCGCAAGGAGATGTCCTCAACAAGAGCTGTGAT	540					
Db	536 CGAATTTCTTCCCATCTTTATCTTGGCTGCGCAAGGAGATGTCCTCAACAAGAGCTGTGAT	595					


```
OY 241 TGCAGTCAGAGGTTGTAAGTTAGCATAAAGCTCCAGAGATGTGCTCTCTGTCA 300
Db |||||
OY 802 TGCAGTCAGAGGTTGTAAGTTAGCATAAAGCTCCAGAGATGTGCTCTCTGTCA 861
Db |||||
OY 301 GACTGTTTTCTCACTGTAAGTTCTGGGTAACTGAGAAAGAGCTTCAACTCTGTCACTG 360
Db |||||
OY 862 GACTGTTTTCTCACTGTAAGTTCTGGGTAACTGAGAAAGAGCTTCAACTCTGTCACTG 921
OY 361 CTTCAGAGTGGGTTTGTGAGATCTCTGCTGTTTCCCTGCTCTGTGAAGAAATCC 420
Db |||||
OY 922 CTTCG----- 926
OY 421 ACTCTAGTCCCTACCTGCAATTTCTCAAGCTTGTCTTACCTGTTCAGCAATTTGGCCAAAC 480
Db |||||
OY 927 ----- 926
OY 481 CGAATTTCTCCAAATCTTTATCTTGCTGCCAGAGATGCTTCACAGAGAGCTGATA 540
Db |||||
OY 927 -----AGAGCTGATG 937
OY 541 CAGCAAGATGGATGTTGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 600
Db |||||
OY 938 CAGCAAGATGGATGTTGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT 997
OY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 660
Db |||||
OY 998 ATCCCGAGTCTCATTTCTGCTGCTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 1057
OY 661 CCGTGGTTGACAAATCAGTAATTTCAATTGAGAAAGCAAAAGCTTCATGATGTTT 720
Db |||||
OY 1058 CCGTGGTTGACAAATCAGTAATTTCAATTGAGAAAGCAAAAGCTTCATGATGTTT 1117
OY 721 CTAGTCACTGTTAGCTGGGATCTCCGCTCCGCCACATGCTATGCTTACATCANG 780
Db |||||
OY 1118 CTAGTCACTGTTAGCTGGGATCTCCGCTCCGCCACATGCTTATGCTTACATCANG 1177
OY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT 840
Db |||||
OY 1178 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACTTACT 1237
OY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGACTATGAGAGAGATTAAAGAC 900
Db |||||
OY 1238 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGACTATGAGAGAGATTAAAGAC 906
OY 901 CAGACT 906
Db |||||
OY 1298 CAGACT 1303
```

Search completed: June 22, 2004, 03:25:21
Job time : 543.914 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 17:08:00 ; Search time 74.9535 Seconds
(without alignments)
6707.969 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_1443
Perfect score: 906
Sequence: 1 atggcccaatgagatgatg9.....agaagataagaaccagact 906

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCBUS.COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	902.8	99.6	1998	4	US-09-816-494-3
2	902.8	99.6	3544	4	US-09-816-494-1
3	365.8	40.4	2377	4	US-09-920-668-3
4	223	24.6	279	4	US-09-016-434-91
5	124.6	11.8	1830	4	US-09-557-921-1
6	106.8	11.3	2303	4	US-09-922-146-3
7	96.2	10.6	2109	4	US-09-016-434-1135
8	96.2	10.6	2109	4	US-09-023-655-946
9	94.6	10.4	1208	4	US-09-023-655-347
10	91.6	10.1	1619	4	US-09-702-705-801
11	91.6	10.1	1619	4	US-09-736-457-801
12	91.6	10.1	1619	4	US-09-614-1248-801
13	91.6	10.1	1619	4	US-09-614-1248-801
14	91.6	10.1	1619	4	US-09-589-184-801
15	91.6	10.1	4637	4	US-09-702-705-804
16	91.6	10.1	4637	4	US-09-736-457-804
17	91.6	10.1	4637	4	US-09-614-1248-804
18	91.6	10.1	4637	4	US-09-614-1248-804
19	91.6	10.1	4637	4	US-09-589-184-804
20	90	9.9	1238	2	US-08-530-290-11
21	90	9.9	1238	4	US-09-702-705-803
22	90	9.9	1238	4	US-09-736-457-803
23	90	9.9	1238	4	US-09-614-1248-803
24	90	9.9	1238	4	US-09-614-1248-803
25	90	9.9	1238	4	US-09-589-184-803
26	90	9.9	2064	4	US-09-702-705-825
27	90	9.9	2064	4	US-09-736-457-825

28	90	9.9	2064	4	US-09-614-1248-825	Sequence 825, App
29	90	9.9	2064	4	US-09-671-325-825	Sequence 825, App
30	90	9.9	2064	4	US-09-589-184-825	Sequence 825, App
31	90	9.9	2109	4	US-09-702-705-826	Sequence 826, App
32	90	9.9	2109	4	US-09-736-457-826	Sequence 826, App
33	90	9.9	2109	4	US-09-614-1248-826	Sequence 826, App
34	90	9.9	2109	4	US-09-671-325-826	Sequence 826, App
35	90	9.9	2109	4	US-09-589-184-826	Sequence 826, App
36	90	9.9	2240	4	US-09-016-434-1100	Sequence 1100, App
37	87	9.6	2000	4	US-09-016-434-1291	Sequence 1291, App
38	86.2	9.5	240	4	US-09-016-434-776	Sequence 776, App
39	80.4	8.9	944	4	US-09-371-671B-10	Sequence 10, App1
40	80.4	8.9	1987	2	US-08-990-379-1	Sequence 1, App1
41	80.4	8.9	1993	2	US-08-990-379-2	Sequence 2, App1
42	79	8.7	539	4	US-09-389-681-311	Sequence 311, App
43	79	8.7	539	4	US-09-620-405B-311	Sequence 311, App
44	79	8.7	539	4	US-09-339-338-311	Sequence 311, App
45	79	8.7	539	4	US-09-433-826B-311	Sequence 311, App

ALIGNMENTS

RESULT 1									
US-09-816-494-3									
Sequence 3, Application US/09816494									
Patent No. 6664085									
GENERAL INFORMATION:									
APPLICANT: Meyers, Rachel A.									
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY									
TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR									
FILE REFERENCE: 10448-030002 US/09/816,494									
CURRENT APPLICATION NUMBER: US/09/816,494									
CURRENT FILING DATE: 2001-03-23									
PRIOR APPLICATION NUMBER: US 60/191,858									
PRIOR FILING DATE: 2000-03-24									
NUMBER OF SEQ ID NOS: 10									
SOFTWARE: FastSeq for Windows Version 4.0									
SEQ ID NO 3									
LENGTH: 1998									
TYPE: DNA									
ORGANISM: Homo sapiens									
US-09-816-494-3									
Query Match 99.6%; Score 902.8; DB 4; Length 1998;									
Best Local Similarity 99.8%; Pred. No. 2.1e-309;									
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	1	ATGGCCCATGAGTGAATGGAATCTCAATTTGTAAGAGGTTGGTCTGCTGGAA	60						
DB	1	ATGGCCCATGAGTGAATGGAATCTCAATTTGTAAGAGGTTGGTCTGCTGGAA	60						
QY	61	AGTGAACGGAAGAGTGTCTAATTTGATAGCGGCAATTTGTGAATCAATATCC	120						
DB	61	AGTGAACGGAAGAGTGTCTAATTTGATAGCGGCAATTTGTGAATCAATATCC	120						
QY	121	CACATTTGGAAGCCATTAATATCAATGCTCCAGACTTGAAGCAAGTTGCAACAG	180						
DB	121	CACATTTGGAAGCCATTAATATCAATGCTCCAGACTTGAAGCAAGTTGCAACAG	180						
QY	181	GACAAAGTGTATTAACAGAGCTCATCGCATTCAGCGAACAATAGTTGACATTGAT	240						
DB	181	GACAAAGTGTATTAACAGAGCTCATCGCATTCAGCGAACAATAGTTGACATTGAT	240						
QY	241	TGCAAGTCAAGAGTTGATTTACGATCAAAAGCTCCCAAGATGTCCTCTCTTCA	300						
DB	241	TGCAAGTCAAGAGTTGATTTACGATCAAAAGCTCCCAAGATGTCCTCTCTTCA	300						
QY	301	GATGTTTCTCTACTGTAATCTGAGTAAGCAAGCAAGCTTCAACTCTGTTCACTG	360						
DB	301	GATGTTTCTCTACTGTAATCTGAGTAAGCAAGCAAGCTTCAACTCTGTTCACTG	360						
QY	361	CTTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	420						

Db 361 CTTCAGAGTGGTTTCTGAGTTCTCTGTTTCTCCGCGCTCTGAGAGAAATCC 420
Qy 421 ACTCTAGTCCCTTACTGCAATTTCTCAGCCTTCTTACCGTTGCAACATTTGGCCAAAC 480
Db 421 ACTCTAGTCCCTTACTGCAATTTCTCAGCCTTCTTACCGTTGCAACATTTGGCCAAAC 480
Qy 481 CGAATTTCTCCAAATCTTATCTTCTGCTGCGAGAGATGCTCTCAACAGAGCTGATA 540
Db 481 CGAATTTCTCCAAATCTTATCTTCTGCTGCGAGAGATGCTCTCAACAGAGCTGATA 540
Qy 541 CAGCAAGATGGATTTGTTATGTTTAAATGCAAGTATACCTGTCCAAAGCCTGACTTT 600
Db 541 CAGCAAGATGGATTTGTTATGTTTAAATGCAAGTATACCTGTCCAAAGCCTGACTTT 600
Qy 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 CCGTGGTTGGCAAAATGATGATTTTCAATTTGAGAAAGCAAAAGCCTCCAAATGATGTT 720
Db 661 CCGTGGTTGGCAAAATGATGATTTTCAATTTGAGAAAGCAAAAGCCTCCAAATGATGTT 720
Qy 721 CTATGCACTGTTTAACTGGGATCTCCGCTCGGCAACATGCTTATGCTTACATCATG 780
Db 721 CTATGCACTGTTTAACTGGGATCTCCGCTCGGCAACATGCTTATGCTTACATCATG 780
Qy 781 AAGAGATGACATGCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 781 AAGAGATGACATGCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Qy 841 ATATCTCCAACTTCAATTTCTGCGGCACTCCGCACTATGAGAAAGAAATTAAGAAC 900
Db 841 ATATCTCCAACTTCAATTTCTGCGGCACTCCGCACTATGAGAAAGAAATTAAGAAC 900
Qy 901 CAGACT 906
Db 901 CAGACT 906

RESULT 2

US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: PHOSPHATASE MOLECULES AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589) ... (2583)
US-09-816-494-1

Query Match 99.6%; Score 902.8; DB 4; Length 3544;
Best Local Similarity 99.8%; Pred. No. 3,3e-309;
Matches 904; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATGGAAGTCAATTTTACTGAGAGTTGGTCTGCTGGA 60
Db 589 ATGGCCCATGAGATGATGGAAGTCAATTTTACTGAGAGTTGGTCTGCTGGA 60
Qy 61 AATGGAACGAAAAAGTGTCTGCTAATTTAGTACCGGCAATTTTGGATATACATATCC 120

Db 649 AGTGAACGAAAAAGTGTCTGCTAATTTAGTACCGGCAATTTTGGATATACATATCC 708
Qy 121 CACATTTTGGAAACCATTAATATCAAGTCTCCAAAGCTTATGAAAGCGAGTTGCAAC 180
Db 709 CACATTTTGGAAACCATTAATATCAAGTCTCCAAAGCTTATGAAAGCGAGTTGCAAC 180
Qy 181 GACAAAGTGTAAATTAACAGCTCATTCAGATTCAGGAAACATAGTTGACATTTGAT 240
Db 769 GACAAAGTGTAAATTAACAGCTCATTCAGATTCAGGAAACATAGTTGACATTTGAT 240
Qy 241 TGCATGCAAGTGTGATTTAAGATCAAAAGCTCCCAAGATGCTCTCTCTTCA 300
Db 829 TGCATGCAAGTGTGATTTAAGATCAAAAGCTCCCAAGATGCTCTCTCTTCA 300
Qy 301 GACTGTTTCTCACTGATCTTGGGTAACCTGAGAAAGCTTCAACTCTGACCTG 360
Db 889 GACTGTTTCTCACTGATCTTGGGTAACCTGAGAAAGCTTCAACTCTGACCTG 360
Qy 361 CTTCAGAGTGGTTTGTGATGTTCTGCTGTTTCTCCGCGCTCTGAGAGAAATCC 420
Db 949 CTTCAGAGTGGTTTGTGATGTTCTGCTGTTTCTCCGCGCTCTGAGAGAAATCC 420
Qy 421 ACTCTAGTCCCTTACTGCAATTTCTCAGCCTTCTTACCGTTGCAACATTTGGCCAAAC 480
Db 1009 ACTCTAGTCCCTTACTGCAATTTCTCAGCCTTCTTACCGTTGCAACATTTGGCCAAAC 480
Qy 481 CGAATTTCTCCAAATCTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 1069 CGAATTTCTCCAAATCTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 541 CAGCAAGATGGATTTGTTATGTTTAAATGCAAGTATACCTGTCCAAAGCCTGACTTT 600
Db 1129 CAGCAAGATGGATTTGTTATGTTTAAATGCAAGTATACCTGTCCAAAGCCTGACTTT 600
Qy 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 1189 ATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 CCGTGGTTGGCAAAATGATGATTTTCAATTTGAGAAAGCAAAAGCCTCCAAATGATGTT 720
Db 1249 CCGTGGTTGGCAAAATGATGATTTTCAATTTGAGAAAGCAAAAGCCTCCAAATGATGTT 720
Qy 721 CTATGCACTGTTTAACTGGGATCTCCGCTCGGCAACATGCTTATGCTTACATCATG 780
Db 1309 CTATGCACTGTTTAACTGGGATCTCCGCTCGGCAACATGCTTATGCTTACATCATG 780
Qy 781 AAGAGATGACATGCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 1369 AAGAGATGACATGCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Qy 841 ATATCTCCAACTTCAATTTCTGCGGCACTCCGCACTATGAGAAAGAAATTAAGAAC 900
Db 1429 ATATCTCCAACTTCAATTTCTGCGGCACTCCGCACTATGAGAAAGAAATTAAGAAC 900
Qy 901 CAGACT 906
Db 1489 CAGACT 1494

RESULT 3

US-09-920-668-3
; Sequence 3, Application US/09920668
; Patent No. 6482644
; GENERAL INFORMATION:
; APPLICANT: Brett P. Cowart
; APPLICANT: Lex M. Cowart
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 8 EXPRESSION
; FILE REFERENCE: RTS-0246
; CURRENT APPLICATION NUMBER: US/09/920,668
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3

LENGTH: 2377
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (135)...(2012)
US-09-920-668-3

Query Match 40.4%; Score 365.8; DB 4; Length 2377;
Best Local Similarity 64.5%; Pred. No. 8.8e-119;
Matches 563; Conservative 0; Mismatches 307; Indels 3; Gaps 1;

28 ATTGTTACTGAGAGTTGTGGCTCTGCTGAAAGTGAACGGAAAAAGTGTGCTGCAATT 87
165 ATGGATGCCAAGAACTGCGCAGCTCTGCGGCGGGCGGCGGCGGCGGCTGCTGATC 224
88 GATAGCGCGCATTTTGTGAATACATATCCACATTTTGGAAAGCATTAATATCAAC 147
225 GACAGCGGCTCTTCTGGAATACACAGCTGGCATGTGCTCAGCTCCGTCAACATCTGC 284
148 TGCTCCAACTTATGAAGCGAAGTTGCAACAGCAAAAGTTTAAATTAACAGCTATC 207
285 TGCTCCAACTTGTGAAGCGCGCTGCAAGCGGCAAGGTGACCATTTGCGAGCTTATC 344
208 CAGCATTCAGCAAAACATTAAGTTGACATTGATTGACAGAGGTTGATTTAGAT 267
345 CAGCGGCTGACAGCGCAGCTGAGAGCTACGAGCCACAGACGTGTGCTTATGAC 404
268 CAAAGCTCCCAAGATGTGCTCTCTCTTCAAGCTGTTTCTCACTGTACTTGGGT 327
405 CAGAGCAGCGGGAGCGCAGAGCTGTGCGGAGACAGCTTCTCTCATCTGCTGAGC 464
328 AAATGAGAAAGCTTCAACTCTGTTCACTGTTGCAAGTGGGTTTGTGAGTTCT 387
465 AAGTGACCGCTGCTTCAAGAGTGGCATCTCACTGCGGCTTGTGCGACCTTCTCC 524
388 CGTTGTTTCCCTGGCTCTGTAAGAAATTCACCT---CTAGTCCCTACCTGATTTCT 444
525 TCCTGCTTCCCGGCTCTGAGAGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 584
445 CAGCTTGTACTTCTGTTGCAATTTGGGCAACCCGAAATTTTCCCAATCTTTATCT 504
585 CAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 644
505 GCGTCCAGGAGATGCTCTCAACAGAGAGTGTATACAGCAAAATGGAGTTGTTATG 564
645 GCGTCCAGGAGAGATGCTCTCAACAGAGATGTTATGACGCAAAATGAATAGCTAGTC 704
565 TTAAATGCGAGCTATACCTGCTCAAAAGCTGACTTTATCCCGAGTCTCATTTCTGCGT 624
705 CTCAACCGCAGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 764
625 GTGCTGTGATGACAGCTTTTGTGAAAAATTTTGCCTGTGTGCAAAATAGTATGAT 684
765 GTCCCATCAACGACACTCTGTGAAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 824
685 TTCAATTGAGAAAGCAAGCTTCAATGATGTTTCTAGTCACTGTTTACTGCGATC 744
825 TTCAATGATTAAGCAAGCTTCTCAAGCTGCTCAAGTCTGCTGCTGCTGCTGCTGCTG 884
745 TTCCGCTCGGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 804
885 TCCGCTCTGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
805 GAAGCTTACAGATTTGTGAAAAAGAAAGCACTATATCTCCAACTTCAATTTTCTG 864
945 GAGCGCTACAGGTTCTGTGAAGGACAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 1004
865 GCGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 897
1005 GCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1037

RESULT 4

US-09-016-434-91
Sequence 91, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
FILING DATE: HEREWITH
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEO ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 279 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGFET03
CLONE: 1234795
US-09-016-434-91

Query Match 24.6%; Score 223; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 8.5e-69;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCATGAGATGATGGAATGGAATTTGTTACTGAGAGTTGTGCTGCTGGA 60
DB 25 ATGGCCATGAGATGATGGAATGGAATTTGTTACTGAGAGTTGTGCTGCTGGA 84
QY 61 AGTGAACGAAAAAGTGTGCTTAATTGATAGCGGCGCATTTTGTGAATTAATCATCC 120
DB 85 AGTGAACGAAAAAGTGTGCTTAATTGATAGCGGCGCATTTTGTGAATTAATCATCC 144
QY 121 CACATTTGGAAGCCATTATATCACTGCTCAAGCTTATGAAGCAAGTTGCAACAG 180
DB 145 CACATTTGGAAGCCATTATATCACTGCTCAAGCTTATGAAGCAAGTTGCAACAG 204

RESULT 5
US-09-557-921-1
Sequence 1, Application US/09557921
Patent No. 6551810

GENERAL INFORMATION:
APPLICANT: Lucche, Ralf M.
TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.416
CURRENT APPLICATION NUMBER: US/09/557,921
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1830
TYPE: DNA
ORGANISM: Homo sapiens
US-09-557-921-1

Query Match
Best Local Similarity 13.8%; Score 124.6; DB 4; Length 1830;
Matches 251; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
QY 461 TTGCCAATCTGGGCGCAACCGAATTTCTTCCATTTTATCTGGCTGCCAGGAGATG 520
Db 1036 TCGAGAACGCTGAGCTCAACCCCAATCTTGGCTTCTGCTTGGCAATGAGGAGATG 1095
QY 521 TCTCAACAGAGAGCTGATACAGACAGATGAGATGTTATGTATGTTAATCCAGTATA 580
Db 1096 CTGAGGACCTGACACCATGACCGGCTGACATGGCTAGCTATCAACCTCACACTC 1155
QY 581 CTTGTCGAAGCTGACTT---ATCCCGAGCTTCATTTCTGCTGCTGCTGCTGATG 637
Db 1156 ATCTTCCCTTACCACTATGAGAAAGCTGTTCACTCAACAGCGGCTCCAGCCACTG 1215
QY 638 ACAGCTTTTGTAGAAATTTTGGCCGTGTTGACCAATGATGATTTTATTGAGAAAG 697
Db 1216 ACAGCAACAGAGAGACCTGCGGAGTCTTGAAGAGCTTTTGAAGTTTATTGAGAAAG 1275
QY 698 CAAAGCCTTCAATGATGTGTCTAGTGCATCTTTTACCTGGAGATCTCCGCTCCGCA 757
Db 1276 CTCACAGTGTGGAGAGGCGTTCTCATCTCCAGCTGCGGAGTCTCCGCTCCGCA 1335
QY 758 CCATGCTATGCGCTCATCATATGAGAGATGACATGTCTTATGATGAGTTTACAGAT 817
Db 1336 CCATGTATGCTTACTTATGATGAGACACATCGGATGACATGATGCTTATTAAT 1395
QY 818 TTGTGAAGAAAAAGACTACTATATCTCCAACTTCAATTTTGGGCCCACTCTCG 877
Db 1396 TTGTCAAGGCAAAAGCAGCATTTCTCCCAACCTTAATCTCATGGGAGATTGCTAG 1455
QY 878 ACTATGAGAGAGATTAGAAC 900
Db 1456 AGTTCAGAGAGACTTAACAC 1478

RESULT 6
US-09-922-146-3
Sequence 3, Application US/09922146
Patent No. 6566133
GENERAL INFORMATION:
APPLICANT: Lex M. Cowser
APPLICANT: Brett P. Monia
TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 9 EXPRESSION
FILE REFERENCE: RTS-0252
CURRENT APPLICATION NUMBER: US/09/922,146
CURRENT FILING DATE: 2001-08-01
NUMBER OF SEQ ID NOS: 48
SEQ ID NO 3
LENGTH: 2303
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (114)...(1268)
US-09-922-146-3

Query Match
Best Local Similarity 11.8%; Score 106.8; DB 4; Length 2303;
Matches 229; Conservative 0; Mismatches 182; Indels 3; Gaps 1;
QY 475 CCAACCGAATTTCTTCCCAATCTTATCTGCTGCGCAGGAGATGCTTCAACAGAG 534
Db 723 CTGTCCAGATCTCTGCCAATCTTATCTGGGAGTCCCGGATTCGCAATTGGAG 782
QY 535 CTGATACAGAGAAATGGATTTGTATGTAAATGCACTATACCTGTCGAA---G 591
Db 783 AGCTGCCAACTGAGGAGATCCGCTACATCTCAATGCAACCCCAACCTCCAAACTTC 842
QY 592 CCGACTTATCCCGAGTCTATTTCTGCGGTGCTGATGATGACAGCTTTGTAG 651
Db 843 TTGGAAGAAATGAGATTTCTTCACTAAGCAATCCCATCTCCACCATGAGCGCAG 902
QY 652 AATATTTGCGGTGTTGACAAATCATGATTTTATGAGAAAGCAAAAGCTTCCAAAT 711
Db 903 AACCTGCGGCTTCTTCCGAGGCGCATGAGTATGATGAGGCTTGTCCAGAAC 962
QY 712 GGATGTCTTCTAGTCACTGTTTACGTTGAGATCTCCGCTCCGACCATGCTATGCGC 771
Db 963 TCGGAGTGTCTGTCACATGCTTGGCGGAGTCAAGCTTCTGTACCGTCACTGTGGCC 1022
QY 772 TACATGAGAGAGATGAGCATGCTTTTATGATGAGCTTACAGATTTGTGAAGAAAA 831
Db 1023 TACCTATGAGAGATCTCACTCTTCTCAACGATGCTTATGATGCTGTGAAGAGAG 1082
QY 832 AGACCTATATCTCAAACTTCAATTTCTGGGCGCAACCTCGGACTATGAG 885
Db 1083 AAGCTTAAGATCTCCCAACTTCACTTATGAGGAGGAGTGTGAGCTTTGAG 1136

RESULT 7
US-09-016-434-1135
Sequence 1135, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1135:
SEQUENCE CHARACTERISTICS:
LENGTH: 2109 base pairs


```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023.655
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 347:
SEQUENCE CHARACTERISTICS:
LENGTH: 1208 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THIRNOT03
CLONE: 1444245
US-09-023-655-347

```

```

Query Match
Best Local Similarity 10.4%; Score 94.6; DB 4; Length 1208;
Matches 142; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```

```

QY 680 TAGATTTGATGAGAAAGCAAAAGCTTCAATGATGATGTTCTAGAGCACTGTTAGCTG 739
DB 4 TTGAGTTTATGAGAGAGAGCTCAGTCAGTGGAGAGGGGCTTTCATCCCTGCGAGGCTG 63
QY 740 GGATCTCCGCTCCGCCACCATGCTATCCCTTCAATGAGAGAGATGAGATGCTT 799
DB 64 GGGGTGCTCCGCTCCGCCACCATGCTATGCTTACTTGAAGACACTCGGATGACCA 123
QY 800 TAGATGAGCTTACAGATTTTGTGAAAGAAAAAGACTCTATATCTCCAACTTCAAT 859
DB 124 TGACTGATGCTTAAATTTGTCAAGGCAAGACCAATATCTCCAACTTCAAT 183
QY 860 TTCTGGGCCCAACTCTGAGATGATGAGAGAAAGATTAGAAC 900
DB 184 TCATGGGGCAGTGTCTGAGATTGAGAGAAAGACTTAAACAC 224

```

RESULT 10

```

US-09-702-705-801
Sequence 801, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongrong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 801
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapien
US-09-702-705-801

```

```

Query Match
Best Local Similarity 10.1%; Score 91.6; DB 4; Length 1619;
Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

```

```

QY 484 ATTCTTCCCAATCTTATCTTGGCTGCGAGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCTCTCTCTTACCTCGGAGAGCTTACCATCATCTGCTCCGGAGAGCATGCTGAC 379
QY 544 CAGAAATGGAGTTGTTATGTTATGATGAGCAAGCTATACCTGTCAAAAGCTGACTTATC 603
DB 380 GCGCTGAGCATCAAGGCTGTTGATATGTTCTTCCGAGCTGCCAAAC---ACTTTGAA 436
QY 604 CCCGAGTCTATTTCTCGCTGCTGCTGATGATGACGCTTTTGTGAGAAATTTTGGCG 663
DB 437 GACACTATACATGATCAAGATCATCCAGTGAAGATACCAAGGCGGACATCAGCTCC 496
QY 664 TGTGTGACAAATCAGTATTTTCAATGAGAAAGCAAAAGCTCCCATGATGATGTTCTA 723
DB 497 TGTTCATGAGAGCATAGATGATCATGATGCTGAGAGAGCTGCGGCGCTGCTG 556
QY 724 GTGCACTGTTTACCTGGATCTCCGCTCCGCTCCATCATGCTATGCTTACATGATGAG 783
DB 557 GTGCACTGCGAGGCGGCAATCTCGCGTCCGCTCCATCATGCTGCTTACCTGATGATG 616
QY 784 AGATGACATGTTCTTATGATGAGCTTACAGATTTGTGAAAGAAAAAGCTTACTATA 843
DB 617 AAGAAACGGGTGAGCTGAGAGAGGCTTCAAGTTCTTGAAGACAGCGCCGAGCATTTATC 893
QY 844 TCTCCAACTTCAATTTTCTGGGCAACTCTGACATGATGAGAAAGAT 893
DB 677 TCCCCCACTTACGCTTCAATGAGGAGCTGCTGAGTTCAGTCCAGGT 726

```

RESULT 11

```

US-09-736-457-801
Sequence 801, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongrong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 801
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-801

```

```

Query Match
Best Local Similarity 10.1%; Score 91.6; DB 4; Length 1619;
Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

```

```

QY 484 ATTCTTCCCAATCTTATCTTGGCTGCGAGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCTCTCTCTTACCTCGGAGAGCTTACCATCATCTGCTCCGGAGAGCATGCTGAC 379
QY 544 CAGAAATGGAGTTGTTATGTTATGATGAGCAAGCTATACCTGTCAAAAGCTGACTTATC 603
DB 380 GCGCTGAGCATCAAGGCTGTTGATATGTTCTTCCGAGCTGCCAAAC---ACTTTGAA 436

```

QY 604 CCGAGTCTGATTTCCGCGGTGCTGTAATGACAGCTTTTGTGAAAAATTTTCCG 663
DB 437 GGACATATACATGACAAAGTCATCCAGTGGAAATACCAAGGCCGACATCAGCTCC 496
QY 664 TGGTTGACAAATCAGATGATTTTCATTTAGAAAAGCAAAAGCTTCATATGATGTCTA 723
DB 497 TGGTTATGAAAGCATAGATGATCATGATGCCGTGAAGATGCGCGTGGCGGTGCTG 556
QY 724 GTGCACTGTTTACGTGGATCTCCGCTCCGCCACATGCTATGCTTACATCATGAAG 783
DB 557 GTGCACTGCGAGGGGCGGATCTCGCGTCCGCCACATCTGCGCTGCTCATGATGATG 616
QY 784 AGATGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAACCTACTATA 843
DB 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTGAGTTCTTTAAGCAGCGCCGACGATTAATC 676
QY 844 TCTCCAACTTCAATTTTCTGCGCAACTCTGAGCTATGAGAAAGAT 893
DB 677 TCGCCCACTTCAAGCTTCATGGGGCAGCTGTGAGTTGAGTCCAGGT 726

RESULT 12
US-09-614-124B-801
; Sequence 801, Application US/09614124B
; Patent No. 6630574

; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-614-124B-801

Query Match 10.1%; Score 91.6; DB 4; Length 1619;
Best Local Similarity 53.2%; Pred. No. 1.3e-21;
Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

QY 484 ATTCTTCCCAATCTTATCTTGTGCTGCCAGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTTCTCTACCTCGGAGTGCCTACCATGCTCCGAGAGACATGCTGAGAC 379
QY 544 CAGAAATGGATGCTATGTTTAAATGCAAGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGGCAATCAAGGCTCTGTGAATGTCCTCGGACGCGCAAC--ACTTTGAA 436
QY 604 CCGAGTCTCATTTCTGCGGTGCTGTAATGACAGCTTTTGTGAAAAATTTTCCG 663
DB 437 GGACATATACATGACAAAGTCATCCAGTGGAAATACCAAGGCCGACATCAGCTCC 496
QY 664 TGGTTGACAAATCAGATGATTTTCATTTAGAAAAGCAAAAGCTTCATATGATGTCTA 723
DB 497 TGGTTATGAAAGCATAGATGATCATGATGCCGTGAAGATGCGCGTGGCGGTGCTG 556
QY 724 GTGCACTGTTTACGTGGATCTCCGCTCCGCCACATGCTATGCTTACATCATGAAG 783
DB 557 GTGCACTGCGAGGGGCGGATCTCGCGTCCGCCACATCTGCGCTGCTCATGATGATG 616
QY 784 AGATGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAACCTACTATA 843

DB 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTGAGTTGTTAAGACAGGCCGACGATTAATC 676
QY 844 TCTCCAACTTCAATTTTCTGCGCAACTCTGAGCTATGAGAAAGAT 893
DB 677 TCGCCCACTTCAAGCTTCATGGGGCAGCTGTGAGTTGAGTCCAGGT 726

RESULT 13
US-09-671-325-801
; Sequence 801, Application US/09671325
; Patent No. 6667154

; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-671-325-801

Query Match 10.1%; Score 91.6; DB 4; Length 1619;
Best Local Similarity 53.2%; Pred. No. 1.3e-21;
Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

QY 484 ATTCTTCCCAATCTTATCTTGTGCTGCCAGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTTCTCTACCTCGGAGTGCCTACCATGCTCCGAGAGACATGCTGAGAC 379
QY 544 CAGAAATGGATGCTATGTTTAAATGCAAGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGGCAATCAAGGCTCTGTGAATGTCCTCGGACGCGCAAC--ACTTTGAA 436
QY 604 CCGAGTCTCATTTCTGCGGTGCTGTAATGACAGCTTTTGTGAAAAATTTTCCG 663
DB 437 GGACATATACATGACAAAGTCATCCAGTGGAAATACCAAGGCCGACATCAGCTCC 496
QY 664 TGGTTGACAAATCAGATGATTTTCATTTAGAAAAGCAAAAGCTTCATATGATGTCTA 723
DB 497 TGGTTATGAAAGCATAGATGATCATGATGCCGTGAAGATGCGCGTGGCGGTGCTG 556
QY 724 GTGCACTGTTTACGTGGATCTCCGCTCCGCCACATGCTATGCTTACATCATGAAG 783
DB 557 GTGCACTGCGAGGGGCGGATCTCGCGTCCGCCACATCTGCGCTGCTCATGATGATG 616
QY 784 AGATGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAACCTACTATA 843
DB 617 AAGAAACGGGTGAGGCTGAGAGAGGCTTGAGTTGTTAAGCAGGCCGACGATTAATC 676
QY 844 TCTCCAACTTCAATTTTCTGCGCAACTCTGAGCTATGAGAAAGAT 893
DB 677 TCGCCCACTTCAAGCTTCATGGGGCAGCTGTGAGTTGAGTCCAGGT 726

RESULT 14
US-09-589-184-801
; Sequence 801, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:

```

; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvik, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-589-184-801
```

```

Query Match      10.1%; Score 91.6; DB 4; Length 1619;
Best Local Similarity 53.2%; Pred. No. 1.3e-21;
Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
```

```

QY 484 ATCTTCCCAATCTTATCTTGCTGCCAGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTTCTCTACTCTGAGAGTCCCTACCAAGCTGCCGAGAGACATGCTGAC 379
QY 544 CAGATGGGATGGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGACATCAAGGCTCTGTGAATGTCCTCGGACGTGCCAACC--ACTTTGAA 436
QY 604 CCCAGTCTGATTTCTGCGTGGCTGGAATGACAGCTTTTGTGAAAATTTTGGCG 663
DB 437 GGAACATATCAGTACAGTGCATCCAGTGAAGATACCAAGGCCACATCAGCTCC 496
QY 664 TGGTTGACAAATCAGTATGATTTTCAATGAGAAAGCAAAAGCTCCAAATGATGTTCTA 723
DB 497 TGGTTCAATGAGACCATAGATGATCATGATGCCGTGAAGAGACTGCCGTGGGCGCTG 556
QY 724 GTGCACGTGTTTACCTGAGATCTCCGCTCCGCCACCATGCTATGCTTACATCATAG 783
DB 557 GTGCATCTGCAGCGCGGACATCTCGCGGTGCGCCACCATCTGCTGCTTACCTGATG 616
QY 784 AGATGACATGCTCTTATGATGAGCTTACAGATTTGTGAAAAGAAAAGACTACTATA 843
DB 617 AAGAAACGGGTGAGCTGAGAGAGGCTTCAGATTCGTTAAGCAGCGCCGACGATTATC 676
QY 844 TCTCCAACTTCAATTTTCTGGGCCAATCTCTGACATGAGAAAGAT 893
DB 677 TGCCCACTTCACTTCAATGAGGAGGAGCTGCTGACGTTGAGTCCAGGT 726
```

```

RESULT 15
US-09-702-705-804
; Sequence 804, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvik, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
```

```

; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 804
; LENGTH: 4637
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-804
```

```

Query Match      10.1%; Score 91.6; DB 4; Length 4637;
Best Local Similarity 53.2%; Pred. No. 2.9e-21;
Matches 218; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
```

```

QY 484 ATCTTCCCAATCTTATCTTGCTGCCAGAGATGCTCTCAACAAGAGCTGATACAG 543
DB 320 ATCTTCCCTTCTCTACTCTGAGAGTCCCTACCAAGCTGCCGAGAGACATGCTGAC 379
QY 544 CAGATGGGATGGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTATC 603
DB 380 GCCCTGGACATCAAGGCTCTGTGAATGTCCTCGGACGTGCCAACC--ACTTTGAA 436
QY 604 CCCAGTCTGATTTCTGCGTGGCTGGAATGACAGCTTTTGTGAAAATTTTGGCG 663
DB 437 GGAACATATCAGTACAGTGCATCCAGTGAAGATACCAAGGCCACATCAGCTCC 496
QY 664 TGGTTGACAAATCAGTATGATTTTCAATGAGAAAGCAAAAGCTCCAAATGATGTTCTA 723
DB 497 TGGTTCAATGAGACCATAGATGATCATGATGCCGTGAAGAGACTGCCGTGGGCGCTG 556
QY 724 GTGCACGTGTTTACCTGAGATCTCCGCTCCGCCACCATGCTATGCTTACATCATAG 783
DB 557 GTGCATCTGCAGCGCGGACATCTCGCGGTGCGCCACCATCTGCTGCTTACCTGATG 616
QY 784 AGATGACATGCTCTTATGATGAGCTTACAGATTTGTGAAAAGAAAAGACTACTATA 843
DB 617 AAGAAACGGGTGAGCTGAGAGAGGCTTCAGATTCGTTAAGCAGCGCCGACGATTATC 676
QY 844 TCTCCAACTTCAATTTTCTGGGCCAATCTCTGACATGAGAAAGAT 893
DB 677 TGCCCACTTCACTTCAATGAGGAGGAGCTGCTGACGTTGAGTCCAGGT 726
```

```

Search completed: June 22, 2004, 02:56:09
Job time : 75.9535 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 21:02:42 ; Search time 7805.33 Seconds
(without alignments)
3692.750 Million cell updates/sec

Title: US-10-029-345A-109
Perfect score: 3418
Sequence: 1 MAHEMTGTVTERLVALLE.....LGKVGSSSTRSGSMETLEIVS 665

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xjh
-Q=/cgn2.1/USPTO.spool/US10029345/runat.21062004.164146.29030/app.query.fasta_1.1294
-DB=genbank -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=spct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10029345.OCGN.1.1.6617@runat.21062004.164146.29030 -NCPV=6 -ICPV=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6.
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenBank1.*
1: gb_ba.*
2: gb_hvg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_stc.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pac.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_hvg_hum.*
31: em_hvg_inv.*
32: em_hvg_other.*
33: em_hvg_mus.*
34: em_hvg_pln.*
35: em_hvg_rod.*
36: em_hvg_mam.*
37: em_hvg_vit.*
38: em_sy.*
39: em_hvg_hum.*
40: em_hvg_mus.*
41: em_hvg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3418	100.0	5450	6 AX482439	AX482439 Sequence
2	3418	100.0	5450	6 AX482478	AX482478 Sequence
3	3406	99.6	1998	6 AX260342	AX260342 Sequence
4	3406	99.6	2732	6 AX180875	AX180875 Sequence
5	3406	99.6	3059	6 AX278461	AX278461 Sequence
6	3406	99.6	3104	6 AX405700	AX405700 Sequence
7	3406	99.6	3496	6 AX441210	AX441210 Sequence
8	3406	99.6	3521	6 AB052156	AB052156 Homo sapi
9	3406	99.6	3544	6 AX260340	AX260340 Sequence
10	3406	99.6	3566	6 AF506796	AF506796 Homo sapi
11	3406	99.6	3766	6 AX374994	AX374994 Sequence
12	3406	99.6	4730	6 BD171157	BD171157 Novel gen
13	3406	99.6	4730	6 BD183422	BD183422 Novel gen
14	3406	99.6	4790	6 AB051487	AB051487 Homo sapi
15	3399	99.4	2102	6 AX713989	AX713989 Sequence
16	3399	99.4	2102	6 AK055973	AK055973 Homo sapi
17	3388	99.1	5111	6 AX482372	AX482372 Sequence
18	3379.5	98.9	2071	6 AX921917	AX921917 Sequence
19	3379.5	98.9	3284	6 BC042101	BC042101 Homo sapi
20	3373	98.7	2118	6 AX099933	AX099933 Sequence
21	3358.5	98.3	2200	6 AX921915	AX921915 Sequence
22	3163	92.5	2807	9 BC031643	BC031643 Homo sapi
23	3096.5	90.6	4943	10 BC059232	BC059232 Mus muscu
24	3096.5	90.6	4975	10 BC057321	BC057321 Mus muscu
25	3079.5	90.1	4874	10 AB052157	AB052157 Mus muscu
26	3057.5	89.5	2756	6 AX482444	AX482444 Sequence
27	3048.5	89.2	1935	6 AY038927	AY038927 Homo sapi
28	3048.5	89.2	3332	6 AX441229	AX441229 Sequence
29	3017.5	88.3	4992	10 AF345951	AF345951 Mus muscu
30	2676	78.3	4827	10 AF345952	AF345952 Mus muscu
31	2016	59.0	172206	9 AC007619	AC007619 Homo sapi
32	2016	59.0	188344	2 AC131617	AC131617 Homo sapi
33	1733.5	50.7	201474	2 AC126682	AC126682 Mus muscu
34	1733.5	50.7	236589	2 AC118035	AC118035 Mus muscu
35	1721	50.4	242590	2 AC133722	AC133722 Rattus no
36	1721	50.4	244605	2 AC097818	AC097818 Rattus no
37	1721	50.4	250782	2 AC128093	AC128093 Rattus no
38	1643.5	48.1	1916	6 AX835305	AX835305 Sequence
39	1643.5	48.1	1916	6 AK098310	AK098310 Homo sapi
40	1505.5	44.0	4026	10 AF345953	AF345953 Mus muscu
41	1422.5	41.6	3231	5 BC063941	BC063941 Danio rer
42	1327	38.8	2264	9 BC045110	BC045110 Homo sapi
43	1326	38.8	2377	6 AR255944	AR255944 Sequence
44	1326	38.8	2377	9 HSU27193	U27193 Human prote
45	1304	38.2	2600	10 BC052705	BC052705 Mus muscu

RESULT 1

ALIGNMENTS

AX482439
 LOCUS AX482439 5450 bp DNA
 DEFINITION Sequence 108 from Patent WO20057460.
 ACCESSION AX482439
 VERSION AX482439.1 GI:22316984
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Todderud C.G., Bol D., Finger J., Lee L., Nelson T., Schieven G., Mcatee, P., Mintler, G., Siemers N., Jackson, D.G. and Ramanathan, C.
 TITLE Polynucleotides encoding human phosphatase
 JOURNAL Patent: WO 02057460-A 108 25-JUL-2002;
 SUIB B RISTOL MYERS CO (US)
 FEATURES
 source
 1. 5450
 location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 538. .2535
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD4451.1"
 /db_xref="GI:22316985"
 /translation="MAHEMIGTQIVTERLVALLSEGTKEVLLIDSPRPVYNTSHILE
 AININCSKLMKRLQODKVLITELIHSAGHKVDIOSKVVYVYDDSDVASLSDC
 FLTVLGLKERSFNSVHLNLAGFAEFSRCPGCEKSTIVPTCLSPCLPVANIGPT
 RILPVLVGCORVANKELIQONGIGVYNASTCCPDPIDPESHELRVAVDSFEK
 KRPTSPNPNFGLQLDYEKKIKVGTASGSKLMLHLEKPNRPVAVSFGQKSE
 TPLSPCADNATSEAGORVHPASVPVSVPISLEPSPVOLSGHLISADRLD
 SNKTKRSPSDDISVSASMAWLSHGFSSSEDALEYKPSITLDTGKTKLCOFSPVQZ
 LSEOTPRISPDKEASLIPKQLQTPRPSDSQSKRLHSRVSSTGTORSLPLRHSQS
 VEDNYHTFLFGLSTQOHLTKSAGLGMGMSDILAPOTSPSLTSSWTFATSSSHR
 YSASATYGGASVAYSCSQLPTCGDQVYSVRRCKPRSPRARSMSWESPFKQKPF
 RRSQMFGRSINSENRSRRELKGVSSQSPSGSMETLEVS"

HAPS

ORIGIN

Alignment Scores:

Pred. No.: 1-4e-204 Length: 5450
 Score: 3418.00 Matches: 665
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-10-029-345a-109 (1-665) x AX482439 (1-5450)

QY 1 MetAlaHieGluMecIIeGIyThrGlnIleValThrgluArgLeuValAlaleuIleuGlu 20
 DB 538 ATGGCCCAAGATGATTTGGAACCAATGTTCAGAGAGTTGGTGGCTGCTGGCA 597
 QY 21 SerGIyThrgluValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 40
 DB 538 AGTGAACGGAAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 657
 QY 41 HisIleuGluAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 60
 DB 658 CACATTTTGAAGCATTATATCACTGCTCCAGCTTATGAGGAGGAGGTTGCAACG 717
 QY 61 AspIyValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 80
 DB 718 GACAAAGGTTAATTACAGCTCATCCAGCATTCAGCAAAACATTAAGGTGACATTGAT 777
 QY 81 CysSerGlnIyValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 100
 DB 778 TGCAGTCGAGAGGTTGATTTACGATCAAGCTCCCAAGATTTGCTCTCTCTCTCTCT 837
 QY 101 AspCysPheLeuThrValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 120

DB 838 GACGATTTCTACCTACTCTTCTGGGTAACTGGAGAGACTTCACTGTTCACTG 897
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlySer 140
 DB 898 CTTCAGAGGAGGTTGCTGAGGTTCTCTGTTGTTCCCTGCGCTCTCTGAGGAAATCC 957
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAlaIleuProThr 160
 DB 958 ACTCTAGTCCCTACTGCTGATTTCTCAGCTTCTGCTTCTGCTTCTGCAACATTTGGCCAAAC 1017
 QY 161 ArgIleuProAsnLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 180
 DB 1018 CCAATTTCTCCCAATCTTTATCTTGGCTGCGCAGCGAGATGCTCTCAACAGAGCTGATA 1077
 QY 181 GlnGlnAsnGlyIleGlyIleValIleuAsnAlaSerIleuIleuIleuIleuIleu 200
 DB 1078 CAGCAGATGGGATTGGTTATGCTTAAATGCGAGCTTAACTGTCCAAGCCTGACTTT 1137
 QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlyIleuIleu 220
 DB 1138 ATCCCGAGTCTCATTTCTGCGCTGCTGCTGATGACGCTTTTGTGAGAAATTTTG 1197
 QY 221 ProTrpLeuAspIySerValAspPheIleuIleuIleuIleuIleuIleuIleuIleu 240
 DB 1198 CCGTGGTTGACAAATCAGTAACTTTGATTAAGAAAGCAAAAGCCTCAATGAGATGTT 1257
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleuIleuIleu 260
 DB 1258 CTAGTCACTGTTAGTGGTGAATCCCGCTCCGCACTGCTTACGCTTACATATG 1317
 QY 261 LysArgMetAspMetSerLeuAspGluIleuIleuIleuIleuIleuIleuIleuIleu 280
 DB 1318 AAGAGGATGACATGCTTTAGATTAAGCTTACGATTTGTGAAAGAAAGAAAGAAAGCTTACT 1377
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnIleuLeuAspIyGlyIleuIleuIleu 300
 DB 1378 AATCTTCAAACTTCAATTTCTGCGCACTCTGAGATGAGAAAGAAAGTAAAGAAC 1437
 QY 301 GlnThrGlyAlaSerGlyProIySerIleuIleuIleuIleuIleuIleuIleuIleu 320
 DB 1438 CAGACTGAGCATCAGAGGCCAAAGAGCAACTGAGCTGCTGAGTGAAGCAAT 1497
 QY 321 GluProValProAlaValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 340
 DB 1498 GAACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1557
 QY 341 CysAlaAspSerAlaThrSerGluAlaIleuIleuIleuIleuIleuIleuIleuIleu 360
 DB 1558 TGTGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1617
 QY 361 ProSerValProSerValGlnProSerIleuIleuIleuIleuIleuIleuIleuIleu 380
 DB 1618 CCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1677
 QY 381 SerGlyLeuHisIleuSerAlaAspArgLeuIleuIleuIleuIleuIleuIleuIleu 400
 DB 1678 AGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1737
 QY 401 SerLeuAspIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 420
 DB 1738 TCTCTGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1797
 QY 421 SerSerSerGluAspAlaLeuGluIleuIleuIleuIleuIleuIleuIleuIleuIleu 440
 DB 1798 TCCTCATGAGAAAGATGCTTGAATATCAAAACCTTCACTACTGTGATGGGACCAAC 1857
 QY 441 LysLeuCysGlnPheSerProValGlnIleuSerGlnIleuIleuIleuIleuIleuIleu 460
 DB 1858 AAGCTATGCAAGTTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1917
 QY 461 AspIyGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 480

Db	1918	GATAGAAGAGAGACGACGATCTCCCAAGAAGCTGCAGACCCGACAGCCTTCAGACAGCAG	1917
Qy	481	SeRlybArqLeuH1sSerVa1ArgThSerserSerglyThra1aGlnArqSerLeuLeu	500
Db	1978	AGCAAGCCATTTGCAATTCGGTCAGAAACCGACGAGCTGGCACCGGCCAGAGATCCCTTTTA	2037
Qy	501	SeRProLeuH1sArqSergLySerVa1GlnbApantYrThsThSerserPheLeuPheGly	520
Db	2038	TCCTCACTGCATCGAAGTGGAGGCTGGAGGACATTCACACACAGCTCTCTTTGGCC	2097
Qy	521	LeuSerThrSergInGlnH1sLeuThrLySerserA1aGlyLeuGlyLeuLySgLyTrpHis	540
Db	2098	CTTTCACACACCGACGACCACTCAGCAAGCTCTCTGGCCCTTGAAGGCTGGCAC	2157
Qy	541	SeRapi1eLeuAlaProGlnThSersThrProSerLeuThrSerserTrpYrPheAla	560
Db	2158	TCGGATATCTTGGCCCCCCACAGACCTCTACCCCTTCCCTGACACACAGCTGGTATTTGGCC	2217
Qy	561	ThrcLluserSerH1sPheYrSerserA1sSerA1a1eYrGlyGlySerA1aSerTrpSer	580
Db	2218	ACAAAGTCTCTACACTTCTACTCTGCTCCGACCACTCTACGAGGACAGGCCAGTTACTCT	2277
Qy	581	AlaYrSerCySerserGlnLeuProThrCySgLyAapGlnValYrSerVa1ArqArg	600
Db	2278	GCTTACACCTGCACGCACTGCCCTCCACTTGGCGGAAACAAAGCTTATCTGTGCGAGCGG	2337
Qy	601	GlnLySProSerAaPaRgAlaAaPSeRArqSergserTrpHisGlnGlySerProPheGln	620
Db	2338	CAGAAGCCAACTGACAGAGCTGACTCGCGCGGAGCTGCATGAAGAGAGCCCTTTGAA	2397
Qy	621	LySglnPheLybArqArqSergCyGlnMetClnPheGlyGlySerL1eMetSergLuan	640
Db	2398	AAGCAAGTTTAAACCAAGAACTGCCAAATGGAAATTTGGAGAGACATCATCTCAGAAAC	2457
Qy	641	ArqSerArqGlnGlnLeuGlnLybVa1GlySergInserSerPheSergLySerMetGln	660
Db	2458	AGGTCAAGGAGAGAGCTGGGAAAGTGGGCACTCACTACTTTCGGCGACATGGAA	2517
Qy	661	Ille1leGlnValSer	665
Db	2518	ATCATTTAGGCTCTCC	2532
RESULT 2			
AX482478		5450 bp	DNA
LOCUS	AX482478		linear
DEFINITION	Sequence 147 from Patent WO02057460.		
ACCESSION	AX482478		
VERSION	AX482478.1	GI:22316999	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Todderd,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G., Suchard,S., Banas,D., Bassolino,D., Feder,J., Krystek,S., Mcatee,P., Mntler,G., Siemers,N., Jackson,D.G. and Ramannathan,C. Polymcleotides encoding human phosphatases		
JOURNAL	Patent: WO 02057460-A 147 25-JUL-2002;		
FEATURES	SQUIBB BRISTOL MYERS CO (US)		
source	Location/Qualifiers		
	1..5450		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	538..2535		
	/note="unamed protein product"		
	/codon_start=1		
	/protein_id="CAD44453.1"		
	/db_xref="GI:22317000"		
	/db_xref="REMBEML.CAD44453"		
	/translatability="MAHEMITGATTTTLEVALLESCTREKULLIDSPREVENTSHILE		
	A1NNCSLUMRR1QODKVLTTEL1QSHARKKVIDCSQKVVYVDSSQDVAS1SSDC		

ORIGIN		FLVLYGKLEKRSNSVHLALAGAEFRSPCPGLCEGSLTPVCIISQPLCPVANIIGPT RI,LPNLYGQORVUNKEILIDONGIGLVNASTCEPXDIPSSHFLRVVNSFCBK ILPMDKSDVPFEKAKRANSGCYVHCISAGISPKATLAIYMRMPSLDEARPYKE KRPLISFNENPLSQLDLYEKRTIKVOTGASBKTKLILHEKREMPPLVBSGQKSE TLPSPCADSATSEAGQRPVHAPASVPSVSPSLSEBPLVQALSGILSARLEB SNLKCSFSILDIKESVSYASMAASLHGFSSDEALERYKSTLIDGNLKCQSPVSE LSQKSPETSPDKESASIPKLOJTPARSSQSKSLSVRTSSGTAQGLSLPLRSHPS VENNHYHSPLFGISTSQOHLTKSAGLGLKMKHDLIAPORTSPSLTSVYFATSSHP YSASAIYGSASYSANSQSLPTCCSDQVYVVRROKRDSPADRSMSHSESPPEKOKP RSCOMEFEGSINSKRSBELGNVGSQSSFSGSMETIEVS"			
ALIGNMENT SCORES:					
Pred. No.:	1.4e-204	Length:	5450		
Score:	3418.00	Matches:	665		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-10-029-345A-109 (1-665) x AX482478 (1-5450)					
QY	1	MechAlhIsgIuMecIlegIYthrgInIleValThrgIuArygLeuValAlaleuLeugIu	20		
DB	538	ATGGCCCATGAGATGATTGGAACTCAAAATTGTACTGAGAGGTTGGTGCTGCTGCGAA	597		
QY	21	SerGIyThrgIuLyValLeuLeuIlaSPSerArGProPheValGIuYrAenThSer	40		
DB	598	AGTGAACCGGAAAAAGCGTCTGCTAAATGAATAGCCGCGCATTTGTGGAAATCAAAATCATCC	657		
QY	41	HiAlIeIeugIuAlIeAnIleAnCySerIyLeuMecIySAyGAyLeuGIIn	60		
DB	658	CACATTTTGGAAAGCATTAATATCAACTGCTCCCAAGCTTATGAAGCAAGGTTGCCAACG	717		
QY	61	AspLyValLeuIleThrgIuLeuIlegIhIseSerIalYshIstLyValAspIleAp	80		
DB	718	GACAAAGCTTAAATTACAGAGCTCATTCAGACTTCAGCGAAACATAAGGTTGACATTGAT	777		
QY	81	CysSerGIuLyValValIalYrAPrGInSerSerGIuAryPValAlaseLeuSerSer	100		
DB	778	TGCAGTCAGAAAGGTTGTAGTTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA	837		
QY	101	AspCyPheLeuThrValIleuLeuGIyLyLeuGIuLySerPheAnSerValHileu	120		
DB	838	GACTGTTTTCACCTGACTTCCTGGGTAACCTGAGAAAGCTTCACTGTTCACCTG	897		
QY	121	LeuAlaGIyGIyPheAlaGIuPheSerArGCyPheProGIyLeuCySGIuGIyLysSer	140		
DB	898	CTTCCAGGTGGGTTGCTGAGTTCTCTCGTTGTTCCCTGGCTCTGTGAAGAAATCC	957		
QY	141	ThrLeuValProThrCyBileSerGIuProCyLeuProValAlaAnIlegIyProThr	160		
DB	958	ACTTAGTCCCTACCTGCAATTCTCACGCTTGCTTACTGTGGCCAACTTGGGCAACC	1017		
QY	161	ArgIleLeuProAnLeuYrLeuGIyCyGIuAryAspValLeuAnLyGIuLeuIle	180		
DB	1018	CGAATTTCTCCCAATCTTAACTCTTGCGCTGCCAGAGATGCTCCCAAGAGACTATA	1077		
QY	181	GIuGIuAnGIyIlegIyYrValLeuAnAlaSerYrThrCySPolyProAspPhe	200		
DB	1078	CAGCAAGATGGAGTTGGTATGTGTTAAAGCCAGCTATACCTTCCAAAGCTGACTTT	1137		
QY	201	IleProGIuSerHlAphLeuAryValProValAnApsSerPheCySGIuLyIleu	220		
DB	1138	ATCCCGAGTCTCATTTCTCGCGGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG	1197		
QY	221	ProTrIleuAspLySerValAspPheIlegIuLyAlaLyAlaSerAnGIyCyVal	240		
DB	1198	CCGTGGTTTGGCAAAATAGATTTATTAAGAGAAACAAAGCTCCAAATGATGTT	1257		
QY	241	LeuValHlAcySLeuAlaGIyIleSerArGSerAlaThrIleAlaIleAlaYrIleMec	260		
DB	1258	CTAAGTCACTGTTTACTGGGATTTCCCGCTCCGCAACATCGTATCGCTTACATCATG	1317		

QY 261 LysArgMetAspMetSerLeuaspGluAlaTyrArgPheValIlysgluYlySarGpProThr 280
Db 1318 AAGAGATGAGCAATCTTGTAGATGAGCTTACAGATTGTGAAAGAAAAGAACTTACT 1377
QY 281 IISerProaspPheaspPheLeuGlyGluLeuaspPheGluGluGluGluGluGluGlu 300
Db 1378 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTGAGACTATAGAGAAAGATTAAAGAC 1437
QY 301 GlnThrGlyAlaSerGlyProIlySerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 320
Db 1438 CAGACTGAGCATCAGGAGCCAAAGAGCAACTCAGCTGCTGCACTGGAGAGCAAT 1497
QY 321 GluProValProAlaValSerGluGlyGluIlySerGluThrProLeuSerProPro 340
Db 1498 GAACCTGTCTCTGCTCTCAGAGGGTGGACAGAAAAGGAGAGCCCTCAGTCCACC 1557
QY 341 CysAlaaspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
Db 1558 TGTGCCAATCTCTACTCTCAGAGGCGAGGAGGAGCAAAAGCCGCTGCATCCCGCAGGCTG 1617
QY 361 ProSerValProSerValGlnProSerLeuLeuGluaspSerProLeuValGlnAlaLeu 380
Db 1618 CCGAGGCTGCCCGCAGCGCTGCGCTTGAAGGACAGCCCGCTGGTACAGCGCTC 1677
QY 381 SerGlyLeuHisLeuSerAlaaspArgLeuGluaspSerAsnIlyLeuIlySarGpPhe 400
Db 1678 AGTGGCTGCCACTCTCTCCGAGAGCGCTGAGAGCAAGCAATAGCTCAGGCTTCTTC 1737
QY 401 SerLeuaspIleIlySerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
Db 1738 TCTCTGAGATATCAATTCAGTTTCATATTACGCCAGCATGCGAGCATCTTACATGGCTTC 1797
QY 421 SerSerSerGluaspAlaLeuGluTyrTyrIlySerProSerThrThrLeuaspGlyThrAsn 440
Db 1798 TCCTCATCAGAAAGATGCTTTGGAAATCTACAAACTTCCACTCTGATGGAGCCAC 1857
QY 441 LysLeuGlyGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
Db 1858 AAGCTATGCCAGTCTCTCCCTGTTCAAGAACTATCGAGAGCATCTCCGAAACCAAGTCT 1917
QY 461 AspIlysgluGluAlaSerIleProIlyIlyLeuGlnThrAlaArgProSerAspSerGln 480
Db 1918 GATAGAGAGAGAACCCAGCATCCCAAGAAAGTGCAGACCGCAGGCTTCAAGACGCGAG 1977
QY 481 SerIlyArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db 1978 AGCAGAGGATTCATTCGCTCAGAAACAGCAGCAGTGGCACTCCGAGAGGTTCTTTTA 2037
QY 501 SerProLeuHisArgSerGlySerValGluaspAsnTyrTrpIlySerPheLeuPheGly 520
Db 2038 TCTCCACTGCATCGAAGTGGAGCGTGGAGGACATTTACACACCACTTCTTTTGGC 2097
QY 521 LeuSerThrSerGlnGlnHisLeuThrIlySerAlaGlyLeuGlyLeuIlyGlyTyrPheHis 540
Db 2098 CTTTCCACAGCAGCAGCAGCAGCTTCAAGAGTCTCTGCTGGGCTTAAAGGCTGGCAC 2157
QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrIlyPheAla 560
Db 2158 TCGGATATCTGGCCCCCAAGCCTTACCTTCCCTGACAGCAGCTGGATTTTGGC 2217
QY 561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlyIlySerAlaSerTyrSer 580
Db 2218 ACAGAGTCTCACTTACTCTGCTGCTCAGCACTTACGAGGAGGAGCTCCAGTTACTCT 2277
QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
Db 2278 GCTTACAGCTGAGCCAGCTGCCCACTGGGAGGAGCAAGTCTATTTCTGTGGCAGGCG 2337
QY 601 GlnIlySerSerAspArgAlaaspSerArgArgSerTyrPheIleGluGluSerProPheGlu 620
Db 2338 CAGAGCCAAAGTACAGAGCTGACTCGCGGAGGAGTGGCAAGAGAGAGCCCTTTGAA 2397

QY 621 LysGlnPheIlySarArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
Db 2398 AAGCAGTTTAAAGCAGAGAGCTGCCAAATGGAAATTTGGAGAGACATCATGTCAAGAC 2457
QY 641 ArgSerArgGluGluLeuGlyIlyValAlaGlySerGlnSerSerPheSerGlySerMetGlu 660
Db 2458 AGGTCAAGGAGAGAGCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTGGGAGAGATGAA 2517
QY 661 IleIleGluValSer 665
Db 2518 ATCATTCAGGTCTCC 2532
RESULT 3
AX260342
LOCUS
DEFINITION Sequence 3 from Patent WO0173059.
ACCESSION AX260342
VERSION AX260342.1 GI:16509305
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
1 Meyers, R. A.
AUTHORS
TITLE
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
JOURNAL
Patent: WO 0173059-A 3 04-OCT-2001;
therefor
Milemnum Pharmaceuticals, Inc. (US)
FEATURES
LOCATION/Qualifiers
1..1998
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1,998-204 Length: 1998
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: 6 Gaps: 0
US-10-029-345A-109 (1-665) x AX260342 (1-1998)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 1 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY 21 SerGlyThrGluIlyValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 61 AGTGAACGGAAGAAAGTCTGCTGCTATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 41 HisIleLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
Db 121 CACATTTTGGAGCCCTTAAATATCAAGTCTCAAGCTTAAAGCAGGATGCAAG 180
QY 61 AspIlyValLeuIleThrGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Db 181 GACAAAGTCTTAAATTCAGAGCTCATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAG 240
QY 81 CysSerGlnIlyValValValTyrAspGlnSerSerSerGlnAspValAlaSerLeuSerSer 100
Db 241 TCGAGTCAGAGAGTGTAGTTTACGATCAAAAGCTCCAAATGTGCTCTCTCTTCA 300
QY 101 AspCysPheLeuThrValLeuLeuGlyIlyLeuGluIlySerPheAsnSerValHisLeu 120
Db 301 GACTGTTTTCATCTGATCTTCTGGTAACTGAAAGGCTTCAACTCTGTTCACTG 360
QY 121 LeuAlaGlyIlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGluIlyIlySer 140
Db 361 CTGAGAGTGGGTTTGTGATGATCTCTGTTTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 420

QY 141 ThrLeuValProThrCysIIleSerGlnProCysLeuProValAlaIleuIIleGlyProThr 160
 DB 421 ACTGTAGTCCCTACCTGCAATTTCTCAGCTTGCTTACTGTGCAACATTTGGGCCAAC 480
 QY 161 ArgIIleuProAsnLeuTyrLeuGlyCysGlnArgAspValIleuAsnIleuLeu 180
 DB 481 CGAATTTCTCCCAATCTTTATCTTGCTGTCAGAGAGATGCTCTCAACAGAGAGCTGATG 540
 QY 181 GlnGlnAsnGlyIIleGlyTyrValIleuAsnIleuSerTyrThrCysProIleuProAspPhe 200
 DB 541 CACAGAGATGGGATGGTTATGTGTAAATGACAGCAATACCTGTCCAAAGCCGTGACTTT 600
 QY 201 IIleProGlnSerIleAspLeuArgValProValAsnAspSerPheCysGlyIleu 220
 DB 601 ATCCCGAGCTCAATTCCTGCGGTGCTGTGAATGACAGCTTTGTGAAATTTTGG 660
 QY 221 ProTrpLeuAspIleuSerValIleuPheIIleGlyIleuValIleuAsnIleuSerVal 240
 DB 661 CCGTGTGGTGAACAATAGTATGATTTTCTTGAAGAAAGCAACCTCCATGATGCTT 720
 QY 241 LeuValIleuCysLeuAlaGlyIIleSerArgSerAlaThrIIleAlaIleuAlaTyrIIleMet 260
 DB 721 CTAGTGCACTGTTAGCTGGGATCTCCGCTCCGCCACCATCGCTATCGCTCATCATG 780
 QY 261 LysArgMetAspMetSerIleuAspGluAlaTyrArgPheValIleuGlyIleuArgProThr 280
 DB 781 AAGAGATGAGACATGCTTTTGAATGAACTTACAGATTTGTGAAGAAAGAAACCTCACT 840
 QY 281 IIleSerProAsnPheAsnPheLeuGlyGlnIleuLeuAspTyrGlyIleuValIleuAsn 300
 DB 841 ATATCTCCAAATCTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTAAGAAC 900
 QY 301 GlnThrGlyAlaSerGlyProIleuSerLysLeuIleuLeuIleuIleuGlyIleuProAsn 320
 DB 901 CAGACTGAGCATCAGGGCCAAAGAGCAACTCAAGCTGTGACCTGGAGAGCAAAAT 960
 QY 321 GlnProValProIleuValSerGlyGlyGlnIleuSerGlnIleuProIleuSerProPro 340
 DB 961 GAACCTGTCTCTCTGTCTCAGAGGGTGAACAGAAAGACAGAGCCCTCATCTCAACCC 1020
 QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 DB 1021 TGTGCCGACTCTCTCACTCAGAGGAGCAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1080
 QY 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
 DB 1081 CCCAGCGTCCAGCGCAGCGCGCTGCTTGAAGAGACAGCCCGCTGTACAGCGCTC 1140
 QY 381 SerGlyIleuIleuSerAlaAspArgLeuGlnAspSerAsnIleuIleuArgSerPhe 400
 DB 1141 AGTGGGTGACCTGTCTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCCGTTCTTC 1200
 QY 401 SerLeuAspIleuIleuSerValSerTyrSerAlaSerMetAlaIleuSerLeuIleuGlyPhe 420
 DB 1201 TCTCTGATATCAATATCATATTCATATTCAGCAGCATGCGACATCTTACATGCTTC 1260
 QY 421 SerSerSerGlnAspAlaLeuGlnIleuTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
 DB 1261 TCTCTCATCAGAAATGCTTGTGAATCTCAAACTTCCACTCATCTGATGGAGCAAC 1320
 QY 441 LysLeuCysGlnIleuSerProValGlnIleuSerGlnIleuProGlnIleuThrSerPro 460
 DB 1321 AACCTATGCAAGTTCCTCTGTTCCAGAACTATCGAGAGACATCCCGCAACAGTCTC 1380
 QY 461 AspLysGlyGlnIleuSerIleuProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 DB 1381 GATTAAGAGAGAACCCAGCATCCCAAGAACTGCAAGCCGCGCTTACAGACAGCAG 1440
 QY 481 SerLysArgLeuIleuSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 DB 1441 AGCAAGAGATGATTCGATCGATCAAGACAGAGAGTGCACCCGCCAGAGATCCCTTTTA 1500

QY 501 SerProLeuIleuArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
 DB 1501 TCTCCACTGATCGAAGTGGAGCGTGGAGACATTTACACACAGCTTCTTTCGGC 1560
 QY 521 LeuSerThrSerGlnIleuIleuThrLysSerAlaGlyIleuGlyIleuLysIleuPhe 540
 DB 1561 CTTTCACAGCAGCAGACACCTCAGAAAGTGTGCTGGCCCTGAGCTTAAAGGCTGGCAC 1620
 QY 541 SerAspIleuLeuAlaProGlnIleuThrSerThrProSerLeuThrSerSerTyrPheAla 560
 DB 1621 TCGATATCTTGGCCCCCAGACCTTACCTCCCTTCAACAGCAGCTGTATTTTGGC 1680
 QY 561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleuTyrGlyIleuSerAlaSerTyrSer 580
 DB 1681 ACGAGTCTCTACACTTCTACTCTGCTCAGCATTTACAGAGCAGTCCGATTTCTCT 1740
 QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
 DB 1741 GCCTACAGCTGACGCCAGCTGCCCACTTCCGAGACCAAGCTATCTGTGCGCAGCGG 1800
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrHisGlnIleuSerProPheGln 620
 DB 1801 CAGAACCCAAAGTACAGACTGCTCGCGAGCTGGCATGAAAGAGAGCCCTTTTGA 1860
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGlySerIleuMetSerGluAsn 640
 DB 1861 AAGCACTTTAAACGCAAGAGCTCCCAATGGAATTTGGAGAGACATCATGTACAGAAC 1920
 QY 641 ArgSerArgGlnIleuGlyIleuValGlySerGlnSerSerPheSerGlySerMetGlu 660
 DB 1921 AGGTCACGGGAAGAGCTGGGGAAGTGGGCACTGCACTTCTTTCGGGACAGCATGAA 1980
 QY 661 IleIIleGluValSer 665
 DB 1981 ATCATTTAGATCTCC 1995

RESULT 4
 AX180875 LOCUS AX180875 2732 bp DNA linear PAT 06-AUG-2001
 DEFINITION Sequence 2 from Patent WO0146394.
 ACCESSION AX180875
 VERSION AX180875.1 GI:15132703
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Ploeman, G.D., Martinez, R., Whyte, D., Manning, G., Sudarshan, S.,
 Hill, R.J. and Flanagan, P.
 Mammalian protein phosphatases
 Patent: WO 0146394-A 2 28-JUN-2001;
 Sugen, Inc. (US)
 FEATURES
 source 1..2732
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3 02e-204 Length: 2732
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-345A-109 (1-665) x AX180875 (1-2732)
 QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
 DB 538 ATGGCCCATGAGATGATTTGAATCTCAATTTTACTGAAGAGTTGGTGGCTGTGAGAA 597

[illegible][illegible]

TITLE Identification of a dual specificity phosphatase: dusp-10
JOURNAL Patent: WO 0177340-A, 1 18-Oct-2001;
MERCK PATENT GmbH (DE)
FEATURES
Source Location/Qualifiers
1. 3059
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
127. 2124
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10545.1"
/db_xref="GI:16605916"
/db_xref="REFSEQ:CAD10545"
/translation="MAHEMIGTDIVTERVALLSESTEKYLIDSPPEVNTSHLE
ALININCSKLMKRLQODKVLITELIGHSAKHKVDIDSKVNVYDSDSDVASLSDC
PLTVLGLKLEKSFNSVHLAAGFAEPRFCPLCEBGRSLVPCISQPLCPVANIPT
RLPVLNLCQORDVNLKELMOQNGIGYVNAANTCPKPFIPESHFLRPVNDPCEK
ILPWLKSDVDFIEKAKASNCVHLAAGISRSATLAIYIMKRMMSIDEAVRPFKE
KPTISPNPFLGQILDYERKIKNOGASGRSKLILHEKNEVPVAVSGGQKSE
TPUSPPCADSATSEAGORVHPASVPSVQPSLLEDSPLVQALSGHLSADRLED
SNLKRKSFLLDKSVSYASMAALHGFSSBEDALEYKPSITLIDGTNLCQPSVQE
LSEQTPETSPDKKESIPKLTQAPRSDSQSKLHSVTRSSGTAQGRSLSPHRSQS
VEDNHTSFLFGISTSOHLTKSAGLKGMSHDILAPOSTSLTSSVYATSESHF
VSAATVGGASAYSAYSCSOLPTCGDOVSVRRROKPDRAISRMSHESPEKQFK
RSCOMEFGBSJMSENRREBELAKVSGSSFSQSMETIEYS"
ORIGIN
Alignment Scores:
Pred. No.: 3.53e-204 Length: 3059
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 0
Query Match: 99.65% Indels: 0
Gaps: 0
US-10-029-345A-109 (1-665) x AK278461 (1-3059)
QY 1 MetAlHiEGlUwEtllEGlYThrgInllEvaIThGluArgLeuValAlaLeuLeuGlu 20
DB 127 ATGGCCCATGATGATGATTGGAACCTCAAAATTGTTACTGAGAGTGGCTGCTGCGGAA 186
QY 21 SerGjYThGluYValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 187 AGTGAACGGAAGAGCTGCTTAATTGATAGCGGCGCATTTTGGAAATCAATACATCC 246
QY 41 HisIlELeuGluAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB 247 CACATTTTGAACCCATTAAATCACTGCTCCAACTTATGAAGGAGTTGCAACAG 306
QY 61 AspLyValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 307 GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAACATAGAGTTGACATTGAT 366
QY 81 CysSerGlnYValYValYValYValYValYValYValYValYValYValYValYValYVal 100
DB 367 TCGAGTCAGAGAGTGTAGTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 426
QY 101 AspCyPheLeuThrValLeuLeuGluYLeuGluYLeuGluYLeuGluYLeuGluYLeuGluY 120
DB 427 GACTGTTTCTCACTGACTCTTCTGGGTAACTGGAGAGAGCTTCAACTCTGTTCACTG 486
QY 121 LeuAlEgLYpheaLagUpheSerArgCysPheProGlyLeuCYsgLugLYpSer 140
DB 487 CTTCGAGAGTGGTGTGCTGAGTCTCTCGTGTCTTCCCTGCTCTGTCGAGGAAATCC 546
QY 141 ThrLeuValProThrCysLeuSerGlnProCysLeuProValAlaLeuLeuLeuLeuLeu 160
DB 547 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGCTTACCTGTTCGCAACATTTGGCCAAAC 606
QY 161 ArgIlELeuProAnbLeuYrLeuGlyCYeGlnArgPValLeuAnbYrGluLeu 180
DB 607 CGAATCTTCCCAATCTTTATCTTGGCTGCGAGAGATGTCTCTCAAGAGAGCTGATG 666

QY 181 GlnGlnAsnGlyIleGlyYrValLeuAnbAlaSerTyThrCysProLeuProAspPhe 200
DB 667 CAGCAAGATGGAGATGGTATGTTAATGACAGCAATACCTGTCCAAAGCTTGACTTT 726
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluYrIleLeu 220
DB 727 ATCCCGAGATCTTCTCTGGGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 786
QY 221 ProThrLeuAspLySerValAspPheIlEGlYValYValYValYValYValYValYVal 240
DB 787 CCGTGGTGTGACAAATCAGTATGATTTCAATGAGAAAGCAAAAGCCCTCAATGATGTGTT 846
QY 241 LeuValAlHiCYLeuAlEgLYIleSerArgSerAlaThrIleAlaIleAlaIleAlaIle 260
DB 847 CTAGTGACATGTTAGCTGGAGATCTCCGCTCCGACCATGCTATCGCTACATCAAG 906
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyArgPheValIleGluYrValYrValYrVal 280
DB 907 AAGAGATGGACATGCTTTAGATGAGCTTACAGATTTGTGAAGAAAAGAGACTACT 966
QY 281 IleSerProAnbPheAnbPheLeuGlyGlnLeuLeuAspTyThrGluYrValYrValYrVal 300
DB 967 ATATCTCCAAATCTCAATTTCTGGGCAACTCTGCACTATGAGAAAGATTAAAGAAC 1026
QY 301 GlnThrGlyAlaSerGlyProLySerTyLeuYrValLeuLeuLeuLeuLeuLeuLeuLeu 320
DB 1027 CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAAGCTCTCCACTGAGAGGCCAAAT 1086
QY 321 GluProValProAlaValSerGluGlyGlnYrSerGlyThrPheLeuSerProPro 340
DB 1087 GAACCTGTCTCTGCTGTCTCAAGGGTGAAGAAAGAGAGAGCCCTCACTGCAACC 1146
QY 341 CysAlaAspSerAlaThrSerGluAlaGlyGlnArgProValHisProAlaSerVal 360
DB 1147 TGTGCGACTGTGTAACCTCAGAGGCGAGGCAAAAGCCCGTGCATCCCGCACGCTG 1206
QY 361 ProSerValProSerValGlnProSerLeuLeuGluYrAspSerProLeuValGlnAlaLeu 380
DB 1207 CCAGGCTCCCAAGCGCTGAGCGCTGTTAGAGAGAGCCCGTGGTACAGGGCTC 1266
QY 381 SerGjYLeuHiELeuSerAlaAspArgLeuGluYrAspSerAsnYrValYrValYrVal 400
DB 1267 AGTGGCTGACCTGTGCGGACAGACAGCTGGAAGACAGATAGCTTCAAGCTTCTTC 1326
QY 401 SerLeuAspIleYrSerValSerTySerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
DB 1327 TCTCTGATATCAATCAATCAGTTCAATTCAGCCAGCATGCGCATCTTACATGCTTC 1386
QY 421 SerSerSerGluAspAlaLeuGlnYrTyTyrProSerThrThrLeuAnbGlyThrAnb 440
DB 1387 TCTCTATCGAAGATGCTTTGGAAATACAAACCTTCACTTCTGATGGAGACCAAC 1446
QY 441 LysLeuCYeGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
DB 1447 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATCGAGACAGACTCCGAAACCACTCT 1506
QY 461 AspLyGluGluAlaSerIleProLyYrLeuGlnThrAlaArgProSerAspSerGln 480
DB 1507 GATAGAGAGAGAGCCAGCATCCCAAGAGAGCTGACAGCCGAGGCTTCAAGACACAG 1566
QY 481 SerLyArgLeuLeuHisSerValArgThrSerSerSerGlyYrThrAlaGlnArgSerLeuLeu 500
DB 1567 ACGAAGCGATTCATTCGCTCGAGAACACAGACAGAGGAGCCGCGAGAGGCTCTTTTA 1626
QY 501 SerProLeuHisArgSerGlySerValGluAspLeuYrHisPheSerPheLeuPheGly 520
DB 1627 TCTTCATCTGATCGAAGTGGAGCTGAGAGACAAATTCACACAGCTTCTCTTTGGG 1686
QY 521 LeuSerThrSerGlnGlnHisLeuThrYrSerAlaGlyLeuGlyLeuYrGlyYrThrHis 540
DB 1687 CTTTCCACAGCAGCAGACAGCACTCAAGAGTGTGCTGCGCTTAAAGGCTGGCAC 1746

QY 541 SerAsp1LeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTyrPheAla 560
 DB 1747 TCGGATATCTTGCCGCCCCAGACCTCTAACCCTCCCTGACCGACGACGTGATATTGGCC 1806
 QY 561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyValYserAlaSerTyrSer 580
 DB 1807 ACAGAGCTCTACACTTCTACTCTGCTCCTCAGCCATCTAGCGAGCGAGTCCAGTACTCT 1866
 QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
 DB 1867 GCCCTACAGCTGACCGACGCTGCCCTGCGAGACCAAGTCTATTCTGTGGCGAGCGG 1926
 QY 601 GlnIleProSerSerAspArgAlaAspSerArgArgSerTrpHisGluGluSerProPheGlu 620
 DB 1927 CAGAGCGCAAGTACAGAGAGCTGCTCGCGCGAGCTGGCATTGAAGAGCCCTTTGAA 1986
 QY 621 LysGlnPheIleValArgTyrSerCysGlnMetGluPheGlyGluSerTlleMetSerGluAsn 640
 DB 1987 AACGAGTTTAAACGCAAGAGCTGCCAATGGAATTGGAGAGACATCATGTCCAGAGAC 2046
 QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 DB 2047 AGGTCACTGGGAAAGACTGGGGAAGTGGGAGTCACTGACTTTCGGGACAGATGGA 2106
 QY 661 IleIleGluValSer 665
 DB 2107 ATCATTTGAGGTCTCC 2121
 RESULT 6
 AX405700 3104 bp DNA linear PAT 14-JUN-2002
 LOCUS Sequence 115 from Patent WO0222660.
 DEFINITION AX405700
 ACCESSION AX405700
 VERSION AX405700.1 GI:21438839
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini, Homiidae; Homo.
 REFERENCE
 AUTHORS Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
 Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.
 TITLE Novel nucleic acids and polypeptides
 JOURNAL Patent: WO 0222660-A 115 21-MAR-2002;
 HYSEQ, INC. (US)
 FEATURES
 source Location/Qualifiers
 1..3104
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 159..2156
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD34807.1"
 /db_xref="GI:21438840"
 /db_xref="EMBL:CAD34807"
 /translation="MAHEMTGQIVPRVALLESTGTEKVLIDSRPFVEVNTSHILE
 AINNSKLMKRRLOODKXVLTRELIGHSAHKVDIDSKRVVYQSSQDVASLSDC
 FLTVLGLKESRPSVTHLAGEFAEFSKCPGLCEKSTLIVPTCSOCLPVANIGPT
 ILIPNLVIGKORVANKELMQNGIGVYANASATCKPPIFESHFLEVPVAFVCK
 ILPMLDSDVPFIKAAASNGCVLVHICAGISRSATLAIAIYIKRMNDLDRAYRCK
 KPTLISFNFNFLOLDYERKIKNOTGAGSPSKLHLHEKNEPVPVAYSEGGKSE
 TPASPCADATSEAGORPVHPASVSVSOPSLILDSPIYQALSGHLSADRLD
 SNLTKRFSLDIKSVSYASMAASLHGFSSSDALFYRPESTLIDGNTKLCOFVDE
 LSEOTBETSPDKEEASIPKKLOTARSPBSQSRKLSVITSSGTAORSLSPLRSGS
 VEDNHTSFLFGLSTGQHLTKSAGIGLGMHSDTLAQTSPSLTSSWYATSSHF
 YSASAIYGGASASASCSQPLTCDQVYSVAROKPDRADSRSMHSESPPEKQK
 RRCQMEPESSTMSERREBELKXVQSFSGSEILIEVS"
 ORIGIN

Alignment Scores:

Pred. No.: 3,61e-204 Length: 3104
 Score: 3406.00 Matches: 663

Percent Similarity: 99.85%
 Best Local Similarity: 99.70%
 Query Match: 99.65%
 DB: 6
 Gaps: 0
 US-10-029-345a-109 (1-665) x AX405700 (1-3104)
 QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnGluValAlaLeuLeuGlu 20
 DB 159 ATGGCCCATGAGATGATGGAACTCAATTCTTACTGAGAGGTGGCTGCTGCGGAA 218
 QY 21 SerGlyThrGlyLysValIleLeuIleLeuAspSerArgProPheValGluTyrAsnThrSer 40
 DB 219 AGTGAAGCGAAGAAAGTGTCTTAATTGATAGCCGGCCATTGTGGAAATACATCATCC 278
 QY 41 HisIleLeuGlnValIleLeuIleLeuAsnCysSerTyrLeuMetLysArgArgLeuGlnGln 60
 DB 279 CACATTTGGAGAGCATTAATATCACTGCTCAAGCTTATGAAAGGATTCGACACAG 338
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
 DB 339 GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 398
 QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB 399 TGCAGTCAGAGAGGTGTAGATTACATCAAGCTCCCAAGATGTGCTCTCTCTTCA 458
 QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeu 120
 DB 459 GACTGTTTCTCACTGACTGCTCTGGTAACTGGAGAGAGAGCTTCACTGTTCACTG 518
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
 DB 519 CTTCGAGAGGTGGTGTGCTGAGCTTCTCTGTTTCCCTGCTGCTGAGAGAAATCC 578
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 579 ACTCTAGCTCTCACTGACTGATTCATTCACCTGCTTCACTGTTGCCAATTTGGCCCAAC 638
 QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuIle 180
 DB 639 CGAATTTCTCCCAATCTTATCTTGGCTGCGACGAGAGTCTCTCAACAAGAGCTGATG 698
 QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
 DB 699 CAGCAAGATGGATGGGTATGTGTTAAATGCCAGCAATACCTGTCAAAGCTGACCTT 758
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
 DB 759 ATCCCGAGTCTCATTTCTGCGTGGCTGCTGAGATGACAGCTTTGTGAGAAAAATTTTG 818
 QY 221 ProTrpLeuAspLysSerValAspPheIleGlyLysAlaLysAsnGlyCysVal 240
 DB 819 CGGTGTTTGGCAAAATCAGTATGATTTCAATTGAGAAAGCAAGCTCCATGATGATGTT 878
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 DB 879 CTAGTGCACTGTTTGTGGGATCTCCCGCTCCGCCACCATCGCTATGCTTACATCAG 938
 QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIlysglyLysArgProThr 280
 DB 939 AAGAGAGATGACAGTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAAGCACTTACT 998
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLysLysIleLysAsn 300
 DB 999 ATATCTCCCACTTCAATTTCTGCGGCAACTCTGACATATAGAGAAAGATTTAAGAAC 1058
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLysLeuLeuHisIleLeuGlyLysProAsn 320
 DB 1059 CAGACTGAGCATATGAGGCCAAAGAGCAAACTCAAGCTCTGCACTGGAGAAAGCAAT 1118
 QY 321 GluProValProAlaValSerGluGlyGlyLysSerGluThrProLeuSerProPro 340

Db 1119 GAACCTGCTCCGTGCTGTCTAGAGGGTGGACAGAAAGGAGAGCCCTCATGTCACCC 1178
 Qy 341 CyeHlaaSpSerAaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 Db 1179 TGTGCGCACTCTGTCTACTCAGAGCGACGAGCAAAAGGCCCGCGCATCCCGCAGGGTG 1238
 Qy 361 ProSerValProSerValGlnProSerLeuLeuGlnAaspSerProLeuValGlnAlaLeu 380
 Db 1239 CCGACGCTGCCCGACGCGTGCAGCCCTCGCTGTAGAGACACCCGCTGGTACAGGCCCTC 1298
 Qy 381 SerGlyLeuHisLeuSerAlaAaspArgLeuGlnAaspSerAnlyLeuValArgSerPhe 400
 Db 1299 AGTGGCTGCAGCTCTCCGACAGACAGCTGGAACAGCATATAGCTCAAGCGTTCTTC 1358
 Qy 401 SerLeuAaspLleuysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 Db 1359 TCTCTGATATCAATATGATTTCAATATTCAGCCGACGTGACGATCTTACATATGCTTC 1418
 Qy 421 SerSerSerGluAaspAlaLeuGluTyrTyrTyrLeuProSerThrThrLeuAaspGlyThrAsn 440
 Db 1419 TCCTCATCAAGATGCTTTGGATTAATACTCAAACTTCCACTCTGGATGGAGCCAC 1478
 Qy 441 LysLeuCysGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGluThrSerPro 460
 Db 1479 AACCTATGCCAGTTCTCCCTGTTCAAGAACTATCGAGAGACCTCCGAAACAGTCTC 1538
 Qy 461 AspLysGlnGlnAlaSerLleProLysLeuGlnThrAlaArgProSerAaspSerGln 480
 Db 1539 GATPAGAGAGAACGACCATCCCGAAGACCTGACGCCCGACCTTACAGACAGCAG 1598
 Qy 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnAaspSerLeu 500
 Db 1599 AGCAAGCATGTCCTTGGTTCAGAACAGACAGCATGTGCACCGCCGACAGTCTTTTA 1658
 Qy 501 SerProLeuHisArgSerGlySerValGluAaspAnlyHisThrSerPheLeuPheGly 520
 Db 1659 TCTCACTGATCGAAGTGGAGGTGGAGGACATTAACACACACACTTCTTTTGGC 1718
 Qy 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrHis 540
 Db 1719 CTTTCCACGACGACGACGACCTTCACAGATCTGTGCGCTGGGCTTAAAGGCTGGCAC 1778
 Qy 541 SerAaspLleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrArgPheAla 560
 Db 1779 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGACAGCTGATTTTGGC 1838
 Qy 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaLleTyrGlyGlySerAlaSerTyrSer 580
 Db 1839 ACAGAGTCCCTCACACTTCTACTCTGCTCAGCCATCTAGGAGGAGGAGTGCAGTTACT 1898
 Qy 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAaspGlnValTyrSerValArgArgArg 600
 Db 1899 GCGTACAGCTGACGACGACCTGACCTTGGGAGACCAAGCTATTTCTGCGCCAGCGG 1958
 Qy 601 GlnLysProSerAaspArgAlaAaspSerArgArgSerThrHisGlnGlnArgProPheGln 620
 Db 1959 CAGAAAGCAAGTACACAGCTGACTCGCGGAGAGCTGGCATAGAGAGACCCCTTTGAA 2018
 Qy 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGlnSerLleMetSerGluAsn 640
 Db 2019 AACAGATTAAAGCAGAACTGCGCAATGGAATTGGAGAGACATCATGTCAGAGAAC 2078
 Qy 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 Db 2079 AGGTTCACGGGAAAGCTGGGGAAGTGGGAGTCAAGTCACTTCTTCCGCGACAGATGAA 2138
 Qy 661 IleIleGlnValSer 665
 Db 2139 ATCATTTAGGTCTCC 2153
 RESULT 7
 AX441210
 LOCUS AX441210 3496 bp DNA linear PAT 28-JUN-2002

DEFINITION Sequence 1 from Patent WO0226997.
 ACCESSION AX441210
 VERSION AX441210.1 GI:21665766
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Lucche, R.M., and Wei, B.
 TITLE Dep-16 dual-specificity phosphatase
 JOURNAL Patent: WO 0226997-A 1 04-APR-2002;
 Ceptyr, Inc. (US)
 FEATURES
 source
 1..3496
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.26e-204 Length: 3496
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: Gaps: 0
 US-10-029-345A-109 (1-665) x AX441210 (1-3496)
 Qy 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 Db 562 ATGGCCATGAGATGATGATGAGTCAATGTTACTAGAGAGTGGTGGCTGCTGGAA 621
 Qy 21 SerGlyThrGlnLysValLeuLeuLleAaspSerArgProPheValGluTyrArgThrSer 40
 Db 622 AGTGAACGGAAGAAATGCTGCTTAATTGATAGCGGCAATTGTGGATACATATCCTC 681
 Qy 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgLeuGlnGln 60
 Db 682 CACATTTTGAAGCCATTATATCACTGCTCCACATTAAGAGGAGGTTGCAACAG 741
 Qy 61 AspLysValLeuLleThrGlnLeuLleGlnHisSerAlaLysHisLysValAaspLleAasp 80
 Db 742 GACAAAGTGTATTAACAGAGCTCATCGACATTCAGCGAAACATAGGTTGACATTGAT 801
 Qy 81 CysSerGlnLysValValTyrAaspGlnSerSerGlnAaspValAlaSerLeuSerSer 100
 Db 802 TGCAGTCAGAGGTTTATGTTATCGATCAAAAGTCCCAAGATGTTCTCTCTTCA 861
 Qy 101 AaspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
 Db 862 GACTGTTTCTCACTTACTTCTGCGGTAACTGGAAGAAAGCTTCAACTCTGTTCACTG 921
 Qy 121 LeuAlaGlyGlyPheHisGlnPheSerArgCysPheProGlyLeuLeuGlyLysSer 140
 Db 922 CTTCAGAGGGGTTTCTAGATCTCTGCTGTTTCTTCCCTGCGCTCTGTGAAGGAAATCC 981
 Qy 141 ThrLeuValProThrCysHisSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 Db 982 ACTCTAGTCCCTTACCTCGCATTTCTCAGCTTGTTTACCTGTTGCCAACAATGGGCCAAC 1041
 Qy 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAaspValLeuAsnLysGlnLeu 180
 Db 1042 CGAATCTTCCCAATCTTATCTTGGCTCCAGCGAGATGCTCTCAACAAAGGAGCTGAAG 1101
 Qy 181 GlnGlnAsnGlyTyrIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAaspPhe 200
 Db 1102 CAGCAAGATGGAGATGGTATGTTAAATGCGAGCAATATCTGTTCCAAAGGCTGACTTT 1161
 Qy 201 IleProGlnSerHisPheLeuArgValProValAsnAaspSerPheCysGlnLysIleLeu 220
 Db 1162 ATCCCGAGTCTCATTTCTGCTGGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 1221


```

OY 221 ProTribLeuAspLysSerValAspPhe11eGluLysAlaLysAlaSerangLysVal 240
DB 1222 CCGTGGTTCGACCAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCATCGATGGATCTGTT 1281
OY 241 LeuValHisCySLeuAlaGly11eSerArgSerAlaThrTleAla11eAlaTyr11eMet 260
DB 1282 CTAGTGCATCTGTTAGCTGGGATCTCCCGCTCCGCCACCAATGCTATCCGCTTACATCATG 1341
OY 261 LysArgMetAspMetSerLeuAspGlu11aTyrArgPheValLysGluLysArgProThr 280
DB 1342 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTGTGAAAGAAAAAGACCTACT 1401
OY 281 11eSerProAsnPheAsnPheLeuGly11eLeuLeuAspTyrGluLysGlu11eLysAsn 300
DB 1402 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAAGAC 1461
OY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHis11eLeuGluLysProAsn 320
DB 1462 CAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTGGAGAGCCAAAT 1521
OY 321 GluProValProAlaValSerGlu11yGly11yLysSerGlu11ThrProLeuSerProPro 340
DB 1522 GAACCTGTCCTGCTGCTCAGAGGGTGGACAGAAAAGCCAGACCCCTCAGTCCACCC 1581
OY 341 CysAlaAspSerAlaThrSerGluAlaAlaGly11aArgProValHisProAlaSerVal 360
DB 1582 TGTGCGACTGCTGCTACTCAGAGGACAGAGCAAGAAAGCCCGTGCATCCCGCCAGCGTG 1641
OY 361 ProSerValProSerValGlnProSerLeuGlu11uAspSerProLeuValGlnAlaLeu 380
DB 1642 CCCACCGTGGCCAGCGTGCAGCGCTCGCTGTAGAGAGACGCCCGTGGTACAGGGCGCTC 1701
OY 381 SerGlyLeuHis11eLeuSerAlaAspArgLeuGlu11uAspSerAsnLysLeuLysArgSerPhe 400
DB 1702 AGTGGGCTGCACCTGTCCTCGAGACAGGCTGGAGAGACAGAAATPAGCTCAAGGCTTCCTTC 1761
OY 401 SerLeuAsp11eLysSerVal11eTyrSerAlaSerMetAla11eSerLeuHis11eGlyPhe 420
DB 1762 TCTCTGATATCAAAATCAGTTTCATATTCAGCCAGATGCGACATCTTACATGCTTC 1821
OY 421 SerSerSerGluAspAlaLeuGlu11yTyrTyrProSerThrThrLeuAspGly11eThrAsn 440
DB 1822 TCCCTCANTCAGAAATGCTTTGGAAATCTCAAACTTCACACTCCTGATGGAGCCACAC 1881
OY 441 LysLeuCySGLnPheSerProValGlnLysLeuSerGlu11uThrProGlu11uThrSerPro 460
DB 1882 AAGCTATGCGCAGTTCTCCCTGTCGAGAACTATGAGACAGACTCCGAAACCAAGTCTT 1941
OY 461 AspLysGluGlu11uAspSer11eProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
DB 1942 GATTAAGAGAGAAAGCCAGCATCCCAAGAGCTGACAGCCGACGCTTCAGACAGCCAG 2001
OY 481 SerLysArgLeuHis11eSerValArgThrSerSerSerGly11uArg11aGluArgSerLeuLeu 500
DB 2002 AACCAAGCCATTCATTCGCTCAGAAACCAAGCAGCATGGCCACCCGCCAGAGTCTCTTTAA 2061
OY 501 SerProLeuHisArgSerGlySerValGluAspAsnTyrHis11eThrSerPheLeuPheGly 520
DB 2062 TCTCAGCTGCATCAGAGTGGAGCGTGGAGAGCAATTAACACACAGCTTCTCTTTCCGCC 2121
OY 521 LeuSerThrSerGlnGlnHis11eLeuThrLysSerAlaGly11eLeuGluLysLeuGly11ePheHis 540
DB 2122 CTTTCCACACAGCCAGACACTCAGCAAGTCTGCTGGCTGGGCTTAAAGGCGCTGGCAC 2181
OY 541 SerAsp11eLeuAlaProGln11uThrSerThrProSerLeuThrSerSerTyrPheAla 560
DB 2182 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCGTGGATTTTCC 2241
OY 561 ThrGluSerSerHisPheTyrSerAlaSerAla11eTyrGly11eSerAlaSerTyrSer 580
DB 2242 ACAGAGCTCCACACTTACTCTGCTCAGCACTTACAGAGGAGGAGTGCAGTACTACT 2301

```

```

OY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnVal11yTyrSerValArgArgArg 600
DB 2302 GCCTACAGCTGCAGCCAGCTGCTCCACTTCCGAGAGCAAGCTATTTCTGTGCGCAGCGCG 2361
OY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrHis11eGluGluSerProPheGly 620
DB 2362 CAGAAAGCCAGTGAACAGAGCTGACTCGCGCGAGCTGCATGAGAGAGAGCCCTTTGAA 2421
OY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGly11uSer11eLeuSerGluAsn 640
DB 2422 AAGCAGTTTAAACCGAGAGCTGCATTAATGAAATTTGAGAGAGCATGTCAGAGAAC 2481
OY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
DB 2482 AGCTACGCGAGAGAGCTGGGAGAAAGTGGCAGTCACTTACTTTCGCGCAGCATGGA 2541
OY 661 11eLeuGluValSer 665
DB 2542 ATCATTTAGAGTCTCC 2556

```

RESULT 8
 AB052156
 LOCUS
 DEFINITION Homo sapiens MKP-7 mRNA for MAPK phosphatase-7, complete cds.
 ACCESSION AB052156
 VERSION AB052156.1 GI:13548676
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Masuda, K., Shima, H., Watanabe, M. and Kikuchi, K.
 MKP-7, a novel mitogen-activated protein kinase phosphatase,
 functions as a shuttle protein
 J. Biol. Chem. 276 (42), 39002-39011 (2001)
 21489429
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLES
 JOURNAL

FEATURES
 source
 1..3521
 Location/Qualifiers
 1..3521
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="12"
 /cell_line="Utrkat"
 1..3521
 /gene="MKP-7"
 564..2561
 /gene="MKP-7"
 /codon_start=1
 /product="MAPK phosphatase-7"
 /protein_id="BAB40814.1"
 /db_xref="GI:13548677"
 /translation="MAHEMTGTQIVTERLVALLESGTEKVLIDSRPFVENVNTHLE
 AININCSKUMRRRLQODKVLITTELIOHSAXHKVDIDCSQKVVYDGSODVASLSDD
 FLTPVLGLKLSFNSVHLAAGFAEFRCFPGCEGKSTIVPTCLISQPCLPVANGIPT
 RILPMLYLGCQDVLDKELKMOQNGIGVLAASNTCKRPDIIPSHFLRVVNDSCFEK
 ILPWLDSKVDPIEKAKASNGCVLVHCLAGISRAITAIATIMKMDMSIDEAYRVCKE
 KRPTSPENFNLGOLLDYERKTIKNOTGASGSKLULHLEKNEPVPVAVSEGCQKSE
 TPLSPCADSATSRNAGQRPVHPASVPSVQPSLIDBSPLVVALSGIHLASDLED
 LSBOTPERTPDKESASTPKRLQTPARPSDSQSKALEHVRTPSSGTQORSLSPHSGS
 VEDNTHYSLFGLSTSQOHLTKSAGLGLKGMHSDILAIPOTSPRLTSSWYPTBSHP
 YSAIYAGSASYSAYSCQLPTCGDQYVSVRRQKPSRRADSRKSRMSWESPEKQK
 RRSQMEFGSISMSNRSRELIKVGSGSPSGMSIIEVS"

ORIGIN

Alignment Scores: 4.3e-204 Length: 3521
 Pred. No.: 3406.00 Matches: 663
 Score: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: 9 Gaps: 0

US-10-029-345A-109 (1-665) x AB052156 (1-3521)

QY 1 MetAlHAGIUNETIEGIYThRGInIleValThRGluARgEuValAlaleuEnglu 20
 DB 564 ATGGCCCATAGATGATTGGAACTCAATTTTACTGAGAGTTGGCTGCTGGAA 623
 QY 21 SerGIYThRGluYsValleuEuiLeaSPSerARgProPheValGIuYrAaenThSer 40
 DB 624 AGTGAACGGAAAGGCTGCTAATTGATAGCCGGCATTTGGAAATACATATCC 683
 QY 41 HisIleleuGIuAlaleuNIleAenCYsSerLYleuMetLYsARgATgLeuGIIn 60
 DB 684 CACATTTTGGAAAGCCATTATATCAACTGCTCCAGCTTATGAAGCGAAGTTGCAACG 743
 QY 61 AspLYsValleuNIleThRGluEuiLeGIInHisSerAlaYsHisLYsValAspIleAaP 80
 DB 744 GACAAAGTGTAAATACAGACTATCCAGCATTCAGCGAAACATAGGTGACATTGAT 803
 QY 81 CySserGIuLYsValValValYrAaPGLInSerSerGIInAspValAlaseRLeuSerSer 100
 DB 804 TGCAGTCAGAAAGTTGATTAGTTAGTTCAGATCAAGCTCCCAAGATTTGCTCTCTTCA 863
 QY 101 AspCYsPheLeuThrValleuLeuGIYsPheGIuLYsSerPheAaenSerValHisLeu 120
 DB 864 GACTGTTTCTCACTGACTTCTGGGTAACTGGAAGAGCTTCACTGTTCACTG 923
 QY 121 LeuAlAGIYGIYpheaAGIuPheSerARgPheProGIYleuCYsGIuGIYLYsSer 140
 DB 924 CTTCAGAGTGGTTTGTCTGAGTCTCTGTTGTTCCCTGCTGTAAGGAAATCC 963
 QY 141 ThrLeuValProThrCYsAlseRGIInProCYsLeuProValAlaAaenIleGIYProThr 160
 DB 984 ACTCTAGTCCCTACCTGCAATTTCTAGCCCTGCTTACCTGTTGCAACATTTGGCCAAAC 1043
 QY 161 ArgIleleuProAaenLeuYrLeuGIYCYsGIInARgAPValleuAaenLYsLeuIle 180
 DB 1044 CGAATTCCTCCCAATCTTATCTTGGCTGCGACGAGATGCTCTCAACAAAGAGCTGATG 1103
 QY 181 GIInGIInAGIYIleGIYrValleuAaenAlaseRrThrCYsProLYsProAaPhe 200
 DB 1104 CAGCAGAAATGGATGGTATGTTAAATGCGACGAAATACCTGCTCAAAAGCTGACTTT 1163
 QY 201 IleProGIuSerHisPheLeuAaenValProValAspSerPheCYsGIuLYsIleLeu 220
 DB 1164 ATCCCGAGTCTCATTTCCGCGCTGTCCTGTAAGACGCTTTGTGGAAGAAATTTTC 1223
 QY 221 ProTrIPLeuAaenLYsSerValAspPheIleGIuLYsAlaLYsAaenAGIYCYsVal 240
 DB 1224 CCTGTGTTGGACAAATCAGTAGATTCTATGAGAAAGCAAAAGCTCCCAATGATGTGTT 1283
 QY 241 LeuValHisCYsValleuAGIYIleSerARgSerAlaThrIleAlaIleAlaYrIleMet 260
 DB 1284 CTAGTCACATGTTAGTGGGATCTCCGCTCCGCAACCATGCTATGCTATCATCATG 1343
 QY 261 LYsARgMetAspMetSerLeuAaenGIuAlaYrARgPheValLYsGIuLYsARgProThr 280
 DB 1344 AAGAGATGAGCAATGCTTTAGATGAAGCTTACAGATTTCTGAAGAAAGAAAGCACTAAT 1403
 QY 281 ILeSerProAaenPheAaenPheLeuGIYGIInLeuAaenPTrGIuLYsIleLYsAaen 300
 DB 1404 ATATCTCCAAACTTCAATTTCTGGGCCCAACTCTGACATATAGAGAAAGATTAAAGAAC 1463
 QY 301 GIInThrGIYAlaseRGIYProLYsSerLYsLeuLYsLeuEuiHisLeuGIuLYsProAaen 320

DB 1464 CAGACTGAGCATTCAGGGCCAAAGACCAATCAAGCTCTGACCTGGAGAACCAAT 1523
 QY 321 GIuProValProAlaValSerGIuGIYGIInLYsSerGIuThrProLeuSerProPro 340
 DB 1524 GAACCTGTCCCTGCTGTCTCAGAGGGTGGACGAAAGAGAGCGCTCCCTGATCCACCC 1583
 QY 341 CYsAlAaSPSerAlaThrSerGIuAlaAGIYGIInARgProValHisPProAlaseRVal 360
 DB 1584 TGTGCGACTGTGCTACCTCAAGAGGACGAGCAAAAGGCCGTGATCCGCCACGTG 1643
 QY 361 ProSerValProSerValGIInProSerLeuLeuGIuAaSPSerProLeuValGIInAlaLeu 380
 DB 1644 CCGAGCTGCCAGCGTGCAGCGCTGCTGTTAGAGGACAGCCCGCTGGATACAGGGCTTC 1703
 QY 381 SerGIYleuHisLeuSerAlaAaPArLeuGIuAaSPSerAaenLYsValYsARgSerPhe 400
 DB 1704 AGTGGGCTCACCTGTCCCGACAGAGCTGGAAGACAGCAATAGCTCAAGCTTCCTTC 1763
 QY 401 SerLeuAaPTrIleYsSerValSerYrSerAlaseRAlaAlaseRLeuHisLYsPhe 420
 DB 1764 TCTCGATATCAAAATCAGTTTCATATTCAGCCAGATGCGACATCTTACATGGCTTC 1823
 QY 421 SerSerSerGIuAaPAlaleuGIuYrYrLYsProSerThrThrLeuAaPGLYrAaen 440
 DB 1824 TCTCATCAAGAGATCCTTTGGAAATCTACAAACCTTCACTGATGGAGACCAAC 1883
 QY 441 LYsLeuCYsGIInPheSerProValGIInLeuSerGIuGIInThrProGIuThrSerPro 460
 DB 1884 AAGCTATGCAATTTCCCTGTTCCAGAACTATCGAAGACAGCTCCCAAAACAGATCCT 1943
 QY 461 AspLYsGIuGIuAlaseRTrIleProLYsLYsLeuGIInThrAlaARgProSerAaSPSerGIIn 480
 DB 1944 GATTAAGAGAGAAAGCCAGATCCCAAGAGCTGCAACCGCAAGCTTTCAGACACCAAG 2003
 QY 481 SerLYsARgLeuHisSerValARgThrSerSerSerGIYThrAlaGIInARgSerLeuLeu 500
 DB 2004 ACGAAGCGATTCATTCGGTCAAGACACAGACAGTGGCACCGCCCAAGAGTCCCTTTTA 2063
 QY 501 SerProLeuHisARgSerGIYSerValGIuAaPArThrHisThrSerPheLeuPheGIY 520
 DB 2064 TCTCCACTCATCAAGATGGAGCGTGGAGGACAAATTAACACACCAAGCTTCCCTTTCCGGC 2123
 QY 521 LeuSerThrSerGIInHisLeuThrLYsSerAlaGIYleuGIYleuLYsGIYTrPHis 540
 DB 2124 CTTCACCAAGCAGAGACCTCAAGAGTGTGCTGGCTTAAGGCTGAGCAC 2183
 QY 541 SerAaPTrIleLeuAlaProGIInThrSerThrProSerLeuThrSerSerTrPYrPheAla 560
 DB 2184 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTAACAGACGCTGATTTTGCC 2243
 QY 561 ThrGIuSerSerHisPheYrSerAlaseRAlaIleYrGIYLYsSerAlaseRTrYrSer 580
 DB 2244 ACAGAGTCTTCACTTACTGCTGCTGACGATCTTACGAGGAGCTCCAGTTACTCT 2303
 QY 581 AlaYrSerCYsSerGIInLeuProThrCYsGIYAaPGLInValYrSerValARgAaAG 600
 DB 2304 GCTTACAGCTGACGACGCTGCCACTTGGAGAGCAAGCTATTTCTGTGCGACAGCGG 2363
 QY 601 GIuLYsProSerAaPArAlaAaSPSerARgAaSPSerTrPHisGIuGIuSerProPheGIu 620
 DB 2364 CAGAGCCAGATGACAGACTACTGCGCGAGCTGCGATGAGAGAGCCCTTTGAA 2423
 QY 621 LYsGIInPheLYsARgARgSerCYsGIInMetGIuPheGIYGIuSerTrIleMetSerGIuAaen 640
 DB 2424 AAGCAGTTTAAACGCAAGAGCTGCAAAATGGAATTTGAGAGACATCATGTCAAGAAC 2483
 QY 641 ARgSerARgGIuGIuLeuGIYLYsValAGIYsSerGIInSerSerPheSerGIYsMetGIu 660
 DB 2484 AGGTCAAGGAGAGAGCTGGGAAAGTGGGCACTGATCTTTCGGGACAGCATGAA 2543
 QY 661 IleIleGIuValSer 665

```

Db      2544 ATCATTGAGTCTCC 2558

RESULT 9
AX260340
LOCUS      AX260340          3544 bp      DNA      linear      PAT 26-OCT-2001
DEFINITION Sequence 1 from Patent WO0173059.
ACCESSION  AX260340
VERSION     AX260340.1  GI:16509303
KEYWORDS
SOURCE
ORGANISM   Homo sapiens (human)
            Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
  1. Meyers, R.A.
    3692 and 21117: dual specificity phosphatase molecules and uses
    therefor
    Patent: WO 0173059-A 1 04-OCT-2001;
    Millennium Pharmaceuticals, Inc. (US)

FEATURES
    source
        1..3544
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD10218.1"
            /db_xref="GI:16509304"
            /translation="MAHEMIGTQIVTERVALLLESTERYVLIDSRPVEYNTSHILE
            AININCKLMKRLQODKVLITELIOHSAKRDIDCSQKVVVYQSSQDVASLESDC
            FLVILGKLEKSFVSLAGAPRSPGGLCBKSKSLVPTCLSQPLPYANIGPT
            ILPLVILGQORDINKEIMQONGITGVILNASNTCPKPIPKSHPIKSHLVVNDSCFK
            KRPTISNPPPLQGLDYEKIKNQTGASGSKATIAIAYIKRMDSLDBAYRPRVE
            TRLSPPCADATSEAAQGRVPASVPSVQPSLIEDSPLYVALSGILHLSARLED
            SNLKRSFSLDIKVSYSASMAASLHGFSSDALRYKPPSLTDGTNKLCSFVDE
            LSEQTPSTFLGKSTSOQHLTKSAGIGLTKGMSDLIAQTSPTSLTSSVFATSSHF
            YSASAIYGGSAVSAYSCSLPCCQQTYSVRKRFSDRADSRMSWHEBSPEFKQPK
            RRSQOMEFGSINSEKRSRELGKVGSSFSFGSMELIEVS"

ORIGIN
Alignment Scores:
Pred. No.:      4.34e-204
Score:          3406.00
Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match:    99.65%
DB:             6
                Gaps:      0

US-10-029-345A-109 (1-665) x AX260340 (1-3544)

QY      1 MetAlahsglumeIleglYThrglnleValThrgluargleuValAlaleuLeuGlu 20
Db      589 ATGGCCCATGAGATGATGGAATCAATTTCTTACTAGAGGTTGGGCTGCTGGGAA 648
QY      21 SerGlyThrgluValleuLeuIlAapSerarGProheValGluTyraenThrser 40
Db      649 AGTGAAGGAAAGAGTGTCTTAATTGATGACCGGCATTGTTGGAATACAAATTCATCC 708
QY      41 HislleuGluAlaIleasnIleasnCySerIleuMeIlyarGargleuGln 60
Db      709 CACATTTTGGAGCCATTATATCACTGCTCCAAAGCTTATGAGGAGGAGGTTGCAACAG 768
QY      61 AspIlyValleuIleThrgluuIleGlnIleSerAlaIlyeHlalyeValAspIleAsp 80
Db      769 GACAAAGTGTATATACAGAGCTCATTCAGCATTCAGCAACATTAAGTTGACATTGAT 828
QY      81 CysSerGlnIlyValValIlyraePglInSerSerGlnAspValAlaSerIleuSerSer 100
Db      829 TGCAGTCAGAGAGGTTGATGTTTACATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 888

QY      101 AspCyPheLeuThrValleuLeuGlyIlyleuGluIlySerPheAsnSerValHisIleu 120
Db      889 GACTGTTTCTCACTGATCTTCTGGGTAACCTGAGAGAGCTTCAACTGCTTCACTG 948
QY      121 LeuAlaGlyGlyPheAlaGluPheSerArgCyPheProGlyIleuCyGluGluIlySer 140
Db      949 CTTCAGAGGTGGGTTTGGAGGTTCTCGTTGTTTCCCTGCTGTGAGAGAAATCC 1008
QY      141 ThrIleValProThrCysIleSerGlnProCysIleuProValAlaasnIleGlyProThr 160
Db      1009 ACTTACGTCCTTACTGATCTTCTCAAGCTTGTCTTACCTGTGCAACATTCGGCCAAAC 1068
QY      161 ArgIleLeuProAsnLeuIlyIleuGlyCyGlnIlyAspValleuAsnIlyleuIle 180
Db      1069 CGAATCTTCCCAATCTTATATCTGGGTGCACAGAGATGCTCTCAACAGAGAGCTGATG 1128
QY      181 GlnGlnAsnGlyIleGlyIlyValIleAsnAlaSerTyThrCysProIlyProAspPhe 200
Db      1129 CAGCAGAAATGGAGTTGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 1188
QY      201 IleProGluSerHisPheIleuArgValProValAsnAspSerPheCyGluIlyIleu 220
Db      1189 ATCCCGAGTCTCATTTCTCGCGTGCCTGTGAATGACAGTTTGTGAGAAATTTTG 1248
QY      221 ProTrpLeuAspIlySerValAspPheIleGluIlyValAlaIlyAsnGlyCyVal 240
Db      1249 CCGTGTGGAGCAAAATCACTAATTTCAATGAGAAAGCAAAAGCTCCATGATGATGTT 1308
QY      241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 260
Db      1309 CTAGTCACCTGTTTGTGCTGGATCTCCGCTCCGCCACCAACCGCTATCCGCTTACATG 1368
QY      261 LysArgMetAspMetSerIleuAspGluAlaTyThrArgPheValIlyGluIlyAspProThr 280
Db      1369 AAGAGAGATGACATGCTTTTGAAGAACTTACAGATTTGTGAAGAAAGAAACCTACT 1428
QY      281 IleSerProAsnPheAsnIleuGlyGlnLeuIleuAspTyIlyGluIlyIlyIleValAsn 300
Db      1429 ATATCTTCAAACTTCAATTTTCTGGGCCCACTCTCGACATATAGAAAGATTAAAGAC 1488
QY      301 GlnThrgIyAlaSerGlyProIlySerIleuIlyleuIleuHisIleuIlySerProAsn 320
Db      1489 CAGACTGAGCATCAGAGCCAAAGAGCAAACTCACTCTCAGCTGAGAGCAAGCAAT 1548
QY      321 GluProValProAlaValSerGluGlyGlnIlySerGluThrProIleuSerProPro 340
Db      1549 GAACCTGTCTCTGCTGTCTCAGAGGGTGGACAGAAAGAGAACCCCTCACTGATCACC 1608
QY      341 CysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSerVal 360
Db      1609 TGTGCCAGCTCTGCTACCTCAGAGGAGAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1668
QY      361 ProSerValProSerValGlnProSerIleuGluIlyAspSerProIleuValGlnAlaLeu 380
Db      1669 CCGAGGCTGCCAGCGCTGAGCGCTGTGTAGAGACAGCCCGCTGTACAGGCCCTC 1728
QY      381 SerGlyLeuHisIleuSerAlaAspArgIleuGluIlySerAsnIlyIlyIlySerPhe 400
Db      1729 AGTGGCTGCACCTGTCCGACAGACAGCTGGAAGACAGCAATTAAGCTCAAGCGTCTTC 1788
QY      401 SerIleuAspIleIlySerValSerTySerAlaSerMetAlaIleSerIleuHisIlyPhe 420
Db      1789 TCTCTGAGATCAATCAATCAATTCATATTCAGCCAGCATGCAAGCTCTTACATGAGCTTC 1848
QY      421 SerSerSerGluAspAlaLeuGluTyThrIlyIlyProSerThrThrIleuAspGlyThraen 440
Db      1849 TCTCATCAGAAAGATGCTTGGATATCTTCAAAACCTTTCATCTCTGATGGGACCAAC 1908
QY      441 LysIleuCyGlnIlySerProValGlnGluIleuSerGluGlnThrProGluIlySerPro 460
Db      1909 AAGCTATGCAAGTCTCCCTGTTTCAAGAACTATCGGAGACAGATCTCCGAAACCAAGTCT 1968

```

Qy 461 AsplyleuGluluaSerileProlyleuGlnThrAlaArgProSerAaspSerGln 480
Db 1969 GATAGAGAGAGAGCAGCATCCCAAGAAAGCTGAGAGCCCGAGCTTACAGACAGCAG 2028
Qy 481 SerIleArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
Db 2029 AGCAAGAGATTGATCGTCAAGAACAGAGAGTGGACCCGCCAGAGTCCCTTTTA 2088
Qy 501 SerProLeuHisArgSerGlySerValGluAAspAntyThiThrSerPheLeuPheGly 520
Db 2089 TCCTCCATCGATCGAAGTGGAGCGTGGAGACAAATTACACACAGCTTCTTTTCGGC 2148
Qy 521 LeuSerThrSerGlnGlnHisleuThrLysSerAlaGlyLeuGlyLeuLysGlyTPHs 540
Db 2149 CTTTCCACCGACCGACGACCTTACGAAATCTGCTGCTGCTGCTTAAAGGCTGGCAC 2208
Qy 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTPYrPheAla 560
Db 2209 TCGGATATCTGGCCCCCGACCTTACCTTCCCTTCCGACGAGCTGTAATTTGCC 2268
Qy 561 ThrGluSerSerHisPheTySerAlaSerAlaIleTyGlyLysSerAlaSerTySer 580
Db 2269 ACAGAGCTCCACACATCTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2328
Qy 581 AlaTySerCysSerGlnLeuProThrCysGlyAAspGlnValTySerValAArgArg 600
Db 2329 GCTTACGCTGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388
Qy 601 GlnLysProSerAspArgAlaAspSerArgArgSerTPHIsGluGluSerProPheGlu 620
Db 2389 CAGAGCGAAGTACAGAGCTGACTGCGCGGAGCTGCGATGAAAGAGAGAGAGAGAG 2448
Qy 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyLysSerIleMetSerGlu 640
Db 2449 AACGAGTTTAAAGCGAAGAGCTGCAATGAAATTTGGAGAGAGCATGATCAGAGAAC 2508
Qy 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db 2509 AGGTACGGAAGAGAGCTGGGGAAGTGGAGATGATGATGATGATGATGATGATGAA 2568
Qy 661 IleIleGluValSer 665
Db 2569 ATCATTGAGGTCTCC 2583
RESULT 10
AF506796 3566 bp mRNA linear PRI 30-OCT-2003
LOCUS Homo sapiens dual specificity phosphatase 16 (DUSP16) mRNA,
DEFINITION complete cds.
ACCESSION AF506796
VERSION AF506796.1 GI:25573087
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 3566)
TITLE Mammalian, Eutherian, Primate, Catarrhini, Homnidae, Homo.
JOURNAL MAPK phosphatase DUSP16/MKP-7, a candidate tumor suppressor for
PUBMED chromosome region 12p12-13, reduces BCR-ABL-induced transformation
1458339 Oncogene 22 (49), 7728-7736 (2003)
REFERENCE
AUTHORS 2 (bases 1 to 3566)
TITLE Hoonbaert, I., Marynen, P. and Baens, M.
JOURNAL Direct Submission
PUBMED Submitted (26-APR-2002) Department for Human Genetics-Flanders
Interuniversity Institute for Biotechnology (VIB), Katholieke
Universiteit Leuven, Herestraat 49, Leuven B-3000, Belgium
FEATURES
Source
1. 3566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

gene
/chromosome="12"
/map="12p13-p12"
/tissue_type="bone marrow"
1. 3566
/gene="DUSP16"
/note="synonym: MKP7"
633. .2630
/gene="DUSP16"
/EC_number="3.1.3.48"
/EC_number="3.1.3.16"
/codon_start=1
/product="dual specificity phosphatase 16"
/protein_id="AA075120.1"
/db_xref="GI:25573087"
/translation="MAHEMTGQIYTERVALLSETEKVLIDSRPFVENVNTHILE
AIVNCSKMKRRLOODKVLITELIOHSAKHKRDIDSCRVVYTDSSQDVASLSSDC
FLTVLGLKLEKFNVAHLAAGPAERFRCFGLCEKSTLVPCLISQCLPVAIVTPT
RILPMLYLCCORDVNLKELMOQNGIGVYVNASNTCPDPFIPESHLRYVDPNFCBK
ILPMLDKSDVFLKAKASNGCVLHCLAGISRSATIAIYIMKRMDSIDEAVRYKE
KRPTISPNFPIGOLIDYERKTKNOTGASGPKSKLHLKRNPEVPVASEGQKSE
TFLSPCADSATSEBAAGRPVYPASVPSVPSVPSVPSVPSVPSVPSVPSVPSVPSVPS
SNLKRSEFLDIKSVYSASMAASLHSESSDRLSRYSTTLDTGRLCQPSVQVE
LSEQPTSPDDEBASIPKQIQTARPSDQSRLHSVRSSTGTRSLSPPLHRSQS
VEDNHTSFLFGLSTSOQHILTKSAGLKGWHSDLIAPOTSPTLSWYFAESSHF
YASNAIYGSAVSAYSQSLPTCGDOVSVRRORPSRDRSRSRHSPEKQFK
RSCQMFPESTMSERSRRELKXVGSQSPFGSMELIIVS"

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Score:
4.38e-204	3566	3406.00
Percent Similarity:	Matches:	663
99.85%	Conservative:	1
Best Local Similarity:	Mismatches:	1
99.70%	Indels:	0
Query Match:	Gaps:	0
99.65%		

DB: 9

US-10-029-345A-109 (1-665) x AF506796 (1-3566)

Qy 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 633 ATGGCCCATGAGATGATGAACTCAATTTGTACGAGAGGTGGTGGCTGCTGCGAA 692
Qy 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyraThrSer 40
Db 693 AGTGAACGGAAGAAAGTCTGCTTAATGATGACCGGCACTTTGTGAAATACATATCC 752
Qy 41 HisIleLeuGluAlaIleAsnIleAsnCySerTyLysLeuMetLysArgArgLeuGln 60
Db 753 CACATTTTGAAGCCATTAAATCACTGCTCCAGACTTATGAGAGGAGGTTGCAACG 812
Qy 61 AsplyleuValleuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIle 80
Db 813 GACAAAGTGTAAATACAGAGCTCACACGATTCAGCAAAATAGAGGTGACATTGAT 872
Qy 81 CysSerGlnLysValValValTyraAspGlnSerSerGlnAspValAlaSerLeuSer 100
Db 873 TGCAGTCAAGAGGTGTAGTTACATCAAGAGCTCCCAAGAGTTCCTCTCTCTTCA 932
Qy 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
Db 933 GACTGTTTCTCACTGCTTCTGCGGTAACGAGAAAGAGCTTCAACTCTGTTCACTG 992
Qy 121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuGlyGluGlyLysSer 140
Db 993 CTTCAGAGGAGGTGGTCTAGTCTCTCGTTTTCCTGCGGCTCTGTAAAGAAATCC 1052
Qy 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 1053 ACTTAGTCTTCACTGCTGCTTCTGAGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1112
Qy 161 ArgIleLeuProAsnLeuTyraLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
Db 1113 CGAATCTTCCCAATCTTATCTTGGCTGCGAGGAGATGTCCTCAAGAGAGTGGANG 1172

QY 181 GlnGluAsnGlyIleGlyTyrValIleuAsnAlaSerTyrThrCysProIysProAspPhe 200
 Db 1173 CAGAGAAATGGATTTGGTTATGGTTAAATGGCAGAAATACCTGCTCAAAAGCTGACCTTT 1232
 QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIleuGlyIleu 220
 Db 1233 ATCCCGAGCTCATTTCTCGGCTGCTCGAATGACAGCTTTTGGAGAAAATTTTG 1292
 QY 221 ProTriLeuAspLysSerValAspPheIleGlyValAlaValAsnGlyCysVal 240
 Db 1293 CCGTGGTTGGACAAATACATGATTTTCATGAGAAAGCAAAAGCTCCAAATGATGTTGTT 1352
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 Db 1353 CTAGTGACCTGTTAGTGGAGTCTCCGCTCCGACCAATCGCTATGCGCTTACATCATG 1412
 QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIlysgIuLysArgProThr 280
 Db 1413 AAGAGATGGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAGACCTACT 1472
 QY 281 IleSerProAsnPheAsnPheLeuGlyValIleuLeuAspTyrGluLysIleLysAsn 300
 Db 1473 ATATCTCAAACTTCAATTTTCTGGGCCAATCCTGACCTATGAGAAAGATTTAAGAAC 1532
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisIleuGluLysProAsn 320
 Db 1533 CAGACTGGAGCATCAGAGGCCCAAGAGCAAACTCAAGCTGCTGACCTGGAGAAAGCCAAAT 1592
 QY 321 GluProValProAlaValSerGluGlyValIleLysSerGluThrProLeuSerProPro 340
 Db 1593 GAACTGTCTGCTGCTCAGAGGGTGGACAAAGAAAGAGAGCGCTCTCACTGCTCACCC 1652
 QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyValArgProValHisProAlaSerVal 360
 Db 1653 TGTCGCCACTGCTGACTTCAGAGGACAGAGCAAGCAAGCCCTGATCCGCCAGCTTG 1712
 QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
 Db 1713 CCCAGCTGCGCCAGCGGTGACCGCTGCTTGAAGAGAGCGCTGCTGACAGCGCTC 1772
 QY 381 SerGlyLeuHisIleuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
 Db 1773 AGTGGCTGACCTGATCCGCAACAGCGCTGAAAGACACCAATAGCTCAACGTTCTTC 1832
 QY 401 SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 Db 1833 TCTCTGATATCAAAATCAGTTTCATATTCAGCCACATGCGCATGCTTACATGCGCTTC 1892
 QY 421 SerSerSerGluAspAlaLeuGluLysTyrLysProSerThrThrLeuAspGlyThrAsn 440
 Db 1893 TCTCTATCAAGAAAGATGCTTTGGAAATACAAACCTTCACTACTCGATGGAGACCAAC 1952
 QY 441 LysLeuCysGlnPheSerProValGlnIleuSerGluGlnThrProGluLysSerPro 460
 Db 1953 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATGAGAGCAAGCTCCGCAACAGCTCT 2012
 QY 461 AspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 Db 2013 GATTAAGAGAAAGCAGCATCCCAAGAAAGCTGACAGCCCGACGCTTCAAGACACCAAG 2072
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 Db 2073 AGCAAGGAGATTGCTTGGCTCAGAACCAAGCAAGAGTGGACCCGCAAGAGTCCCTTTTA 2132
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
 Db 2133 TCTCATCTGATCGAAGATGGAGAGCTGGAGAGCAATTAACAACAGCTCTCTTTTGAGC 2192
 QY 521 LeuSerThrSerGlnGlnIleuThrLysSerAlaGlyLeuGlyLysLysGlyTyrHis 540
 Db 2193 CTTTCCACCAAGCAAGCACTTCAAGAGTGTGCTGCGCTTAAAGGGCTGAGC 2252

QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
 Db 2253 TCGGATATCTTGGCCCGCCCAAGACCTTACCTTCCCTGACAGACATGATTTTGCC 2312
 QY 561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
 Db 2313 ACAGAGTCTTCACTTACTTCTGCTTCAAGCCATTCAGAGAGCAGTGCAGCTTACTCT 2372
 QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
 Db 2373 GCTTACAGCTCCAGCCAGCTGCCACTTGGAGAGACCAAGCTTATCTGCGCAGCGG 2432
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerThrHisGluGluSerProPheGlu 620
 Db 2433 CAGAAAGCAAGTGCACAGCTGACTGCGGCGGAGCTGGCAGATGAAAGAGCCCTTTGAA 2492
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGlySerIleMetSerGluAsn 640
 Db 2493 AAGCAGTTTAAAGCAGAAAGCTCCAAATGGAATTTGGAGAGACATCATGTCAGAGAAC 2552
 QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 Db 2553 AGGTCAAGGAAAGAGCTGGGAGAAAGTGGCAGTCAAGTCTTTCGAGGAGCATGAA 2612
 QY 661 IleIleGluValSer 665
 Db 2613 ATCATTTAGAGTCTCC 2627

RESULT 11
 AX374994
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS
 Tang, Y. T., Elliott, V. S., Ramkumar, J., Yao, M. G., Burford, N.,
 Wang, Y. E., Stewart, E. A., Gandhi, A. R., Patterson, C., Lee, E. A.,
 Hatfield, A. J., Lu, D. A., Tribouley, C. M., Griffin, D. A., Baughn, M. R.,
 Yue, H., Warren, B. A., Nguyen, D. B. and Wallis, N. K.
 Protein phosphatases
 Patent: WO 0210363-A 17 07-FEB-2002;
 Incyte Genomics, Inc. (US)
 Location/Qualifiers
 1..3766
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 7480570CBI"

ORIGIN
 Alignment Scores:
 Pred. No.: 4,72e-204 Length: 3766
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 0
 Query Match: 99.65% Indels: 0
 DB: 6 Gaps: 0

US-10-029-345A-109 (1-665) x AX374994 (1-3766)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
 Db 538 ATGGCCCAAGAGATGATTTGAACTCAAAATGTACTGAGAGCTTGGGCTCTCTCGAA 597
 QY 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 Db 598 AGTGAAGGAGAAAGTGTCTGTAATTAAGTAGCCGGCCATTGTGGAAATCAATATACATCC 657

PD WO 02052005-A/13
 PD 04-JUL-2002
 PF 20-DEC-2001 WO 2001JP011217
 PR 22-DEC-2000 JP 00P 389742
 PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
 PC C12N15/12.C07K14/47
 CC Novel gene and protein encoded thereby
 FH Key Location/Qualifiers
 FT CDS (184) . (2178) .
 Location/Qualifiers
 source 1..4790
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="Laxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6,61e-204 Length: 4790
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 Gaps: 0
 DB: 6
 US-10-029-345A-109 (1-665) x BD171157 (1-4790)
 QY 1 MetaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
 DB 184 ATGCCCATGAGATGATGAGACCAATTGTACTGAGAGGTGGCTCTGCTGGAA 243
 QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGlnIleAsnThrSer 40
 DB 244 AGTGAACGGAAGGAGTGTGTAATGATGAGCGGCCCATTTGTGGAATACAAATGCATCC 303
 QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerValLeuMetLeuArgArgLeuGln 60
 DB 304 CACATTTTGGAGCCATTAATATCACTCTCCAGCTTATGACGGAAGGTGGAAAG 363
 QY 61 AspPheValLeuIleThrGluLeuIleGlnIleSerAlaValHisIleValAspIleAsp 80
 DB 364 GACAAAGTGTATTCAGAGCTCATCCAGCATTCAGGAAACATTAAGTTGACATTGAT 423
 QY 81 CysSerGlnValValValValIleAspIleSerSerGlnAspValAlaSerLeuSerSer 100
 DB 424 TGCAGTCAGAGAGTGTATTCATCATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 483
 QY 101 AspCysPheLeuThrValLeuLeuGlnIleValLeuGlnIleValLeuIleValLeu 120
 DB 484 GACTGTTTCTCAGCTGACTCTCTGCTAACTGGAGAAAGCTTCAACTCTGTTCACTG 543
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyIleSer 140
 DB 544 CTTCGAGGAGGTGTGCTGAGGTTCTCTGTTGTTTCCCTGCTCTGTGAAGAAATCC 603
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 604 ACTTAGTCCCTTACCTGATTTCTCAGCTCTGCTTACTTGTCCAAACATTTGGCCAAAC 663
 QY 161 ArgIleLeuProAsnLeuIleValCysGlnArgAspValLeuAsnIleGluLeuIle 180
 DB 664 CGAATCTCTCCCATCTTATCTTGCTGCTCAGAGAGATGCTCTCAACAAGAGCTGATG 723
 QY 181 GlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerIleThrCysProIleProAspPhe 200
 DB 724 CAGCAGATGGAGTGTATGTTAAATGACAGCAATACCTGTCCAAAGCTGACTTT 783
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIleLeu 220
 DB 784 ATCCCCAGATCTCATTTCTGCTGCGCTGTAATGACAGACTTTTGTGAAAAATTTTG 843
 QY 221 ProIleLeuAspIleSerValAspPheIleGluValAlaValAlaSerAsnGlyCysVal 240
 DB 844 CCTGTGTGACAAATCAATGATTTCAATTTGAGAAAAGCAAAAGCTCCAAATGATGTGTT 903

QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
 DB 904 CTAGTGCACGTGTTAGCTGGAGATCTCCGCTCCGCCACATCGCTATGCTCATCATG 963
 QY 261 LysArgMetAspMetSerLeuAspGluAlaIleArgPheValIleGlyIleValProThr 280
 DB 964 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGCACTACT 1023
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuAspIleGlnIleValIleValAsn 300
 DB 1024 ATATCTCCAACTTCAATTTTCTGGGCCAATCTCTGACTATGAGAAAGATTAAGAAC 1083
 QY 301 GlnThrGlyAlaSerGlyProIleSerValLeuValLeuIleHisIleGlnIleValProAsn 320
 DB 1084 CAGACTGAGATCAGAGGCCAAAGACCAACTACAGCTGTGACCTGAGAGAGCAAT 1143
 QY 321 GluProValProAlaValSerGluGlyGlnIleValSerGluThrProLeuSerProPro 340
 DB 1144 GAACCTGTCTCTCTGCTCAGAGGTTGACAGAAAAGCAAGCGCCCTCATGTCACCC 1203
 QY 341 CysAlaAspSerAlaThrSerGluAlaIleGlyIleValProValHisProAlaSerVal 360
 DB 1204 TGTGCCAGCTCTGCTACTCAGAGCAGAGAGCAAGAGCCCTGTGCATCCGCCAGCTG 1263
 QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
 DB 1264 CCAGAGGTGCCAGCTGAGCGCTGTGTTAGAGGACAGCGCGCTGTACAGGCGCTC 1323
 QY 381 SerGlyLeuHisIleSerValAspArgLeuGluAspSerAsnIleValValArgSerPhe 400
 DB 1324 AGTGGCTCAGCTGTCTGCGAGACAGCTGGAAAGCACAATTAAGCTCAAGCTTCTTC 1383
 QY 401 SerLeuAspIleValSerValSerIleSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 DB 1384 TCTTGAGATCAATCAATGTTTCAATTAACAAGACAGAGATGAGATCTTACATGCTTC 1443
 QY 421 SerSerSerGluAspAlaLeuGlnIleValIleValProSerThrThrLeuAspGlyIleAsn 440
 DB 1444 TCTCATCAGAGATGCTTTGAAATACACAACTTCCACTCTCGATGTGGAGCAAC 1503
 QY 441 LysLeuCysGlnPheSerProValGlnIleLeuSerGluGlnIleThrProGluThrSerPro 460
 DB 1504 AAGCATCCAGGTTCTCCCTGTTCAAGAACTATCGAGAGACTCCGAAACAGTCCCT 1563
 QY 461 AspPheGluGlnAlaSerIleProValValLeuGlnIleThrAlaArgProSerAspSerGln 480
 DB 1564 GATTAAGAGAGAGCCAGATATCCCAAGAGCTCAGACCGCCAGGCTTCAAGACAGCCAG 1623
 QY 481 SerIleArgLeuHisIleValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 DB 1624 AGCAAGGAGTGTGCTGTCAGAACAGAGCAGAGTGACGCCCGCAAGGTCCCTTTTA 1683
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnIleHisThrSerPheLeuPheGly 520
 DB 1684 TCTCATCTGATGAGATGAGAGTGGAGCGTGAGAGCAATTAACAACAGCTTCTTTTCCG 1743
 QY 521 LeuSerThrSerGlnGlnIleLeuThrIleValSerAlaGlyIleGlyIleValIleHis 540
 DB 1744 CTTCACACAGCAGCAGCAGCAGCTCAAGAGTCTGCTGAGCTTGAAGGGCTGGGAC 1803
 QY 541 SerAspIleLeuAlaProGlnIleThrSerThrProSerLeuThrSerSerIlePheAla 560
 DB 1804 TCGGATATCTGGCCCCCAGACCTTCACTTCCCTGACCAAGCAGCTGATATTGTC 1863
 QY 561 ThrGluSerSerHisPheIleSerAlaSerAlaIleIleArgIleValIleSerIleSer 580
 DB 1864 ACAGAGTCTCAGACTTCACTCTGCTCAGGCACTTACGAGAGAGGAGCCAGTTACTCT 1923
 QY 581 AlaIleIleSerCysSerGlnLeuProThrCysGlyIleAspGlnValIleSerValArgArg 600
 DB 1924 GCTTAAGCTGACAGCAGCTGCCACTTGGAGAGCAAGTCTATCTGTGCGGAGGCGG 1983

QY 601 GlnlybProSerAaPArgAlaAaPSeArGArSerTrpHlgluJlusePProPheGlu 620
 DB 1984 CAGAGCCAACTGACAGAGCTGACTCCGCGCGAGCTGCGATGAAGAGCCCTTTGAA 2043
 QY 621 LysGlnPheLysAaGArSerCyGlmMetGluPheGluSerIlleMetSerGluAsn 640
 DB 2044 AAGCAGTTTAAACGCAAGAGCTGCCAATGGAATTTGAGAGACATCATCTCAGAGAAC 2103
 QY 641 ArgSerArgGluGluLysGluLysValGlySerGlnSerPheSerGlySerMetGlu 660
 DB 2104 AGGTCAAGGAGAGAGCTGGGAGAAAGTGAGCAGTCAGTCTTTTGGGCGACGATGAA 2163
 QY 661 IleIleGluValSer 665
 DB 2164 ATCATTTAGGCTCTCC 2178
 RESULT 13
 LOCUS BD183422 4790 bp DNA linear PAT 17-JUN-2003
 DEFINITION Novel genes and proteins encoded by the genes.
 ACCESSION BD183422
 VERSION BD183422.1 GI:31875622
 KEYWORDS JP 2002345492-A/135.
 SOURCE JP 2002345492-A/135.
 ORGANISM Homo sapiens (human)
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 4790)
 AUTHORS Ohara,O., Nagase,T. and Nakajima,D.
 TITLE Novel genes and proteins encoded by the genes
 JOURNAL Patent: JP 2002345492-A 135 03-DEC-2002;
 KAZUSA DNA RESEARCH INSTITUTE
 COMMENT OS Homo sapiens (human)
 FN JP 2002345492-A/135
 PD 03-DEC-2002
 PI 26-FEB-2002 JP 2002049009
 PR OSANU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
 PC C12N15/09,C07K14/47//A61K31/711,A61K38/00,A61K48/00,A61P25/00,
 PC A61P25/14,
 PC A61P25/18,A61P35/00,C12N15/00,A61K37/02
 CC Novel genes and proteins encoded by the genes FH Key
 Location/Qualifiers
 FT CDS (184)..(2178).
 FEATURES
 Source 1..4790
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.61e-204 Length: 4790
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-345a-109 (1-665) x BD183422 (1-4790)
 QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 DB 184 ATGGCCCATAGAGATGGAGTGAAGCTCAAAATGTGTACTAGAGAGTGGGCTGTGGTGA 243
 QY 21 SerGlyThrGluLysValLeuLeuIleAaPSeArGArSerPheValGluIlyrAaThrSer 40
 DB 244 AGTGCAGCGGAAAGAGTGTCTAATTGATAGCCGCGCATTTGGAGAAATCAATATACATCC 303
 QY 41 HisIleLeuGlnValAlaIleAsnIleAaPSeArGArSerIlyLeuMetLysValGluLeuGln 60
 DB 304 CACATTTTGAAGGCAATTAATATCACTGCCAAGCTTATGAAGCGAAGTTGCAACAG 363
 QY 61 AspLysValIleuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAaPSeAr 80

DB 364 GACAAGGTATTATTACAGAGCTCATCCAGCATTTACGCGAAACATAAGTTGACATTGAT 423
 QY 81 CysSerGlnLysValValValIlyrAaPSeArSerGlnAaPValAlaSerLeuSerSer 100
 DB 424 TGCAGTCAGAGAGGTGTAGTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 483
 QY 101 AspCPheLeuThrValLeuLeuGluLysLeuGluLysSerPheAaSerValHisLeu 120
 DB 484 GACTGTTTCTCACTGTAATCTTGAGTAACTGAGAGAGCTTCAACTGTGTACCTG 543
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArGArPheProGlyLeuGluGlyLysSer 140
 DB 544 CTTCGAGGTGGGTTCCTGAGATTCTCTGTTGTTCCCTGCTCTGTGAAGAAATCC 603
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 604 ACTCTAGTCCCTACCTGACATTTCTCAGCCCTTGCTTACCTGTCCAATGGGCGCAACC 663
 QY 161 ArgIleLeuProAsnLeuIlyrLeuGlyCysGlnAaPValLeuAsnLysGluLeuIle 180
 DB 664 CGAATTTCTCCCAATCTTATCTTGCTGCTGACAGAGATGCTCTCAACAGAGCTGATG 723
 QY 181 GlnGlnAsnGlyIleGlyIlyrValLeuAsnAlaSerIlyrThrCysProLysProAaPhe 200
 DB 724 CACAGAGATGGAGTTGGTTATGTGTAAATGCCAGCAATATCTGTCCAAAGCTGACTTT 783
 QY 201 IleProGluSerHisPheLeuArgValProValAsnAaPSeArPheCysGluLysIleLeu 220
 DB 784 ATCCCGAGCTCATTTCCGCTGCTGCTGTAAGACAGCTTTGTGTGAGAAATTTTG 843
 QY 221 ProThrLeuAspLysSerValAaPheIleGluLysValAlaLysSerAsnGlyCysVal 240
 DB 844 CCGTGTGTGACAAATCAGTAGATTTTCATTGAGAAAGCAAAACCTCCCAATGAGATGCTT 903
 QY 241 LeuValHisCysLeuAlaGlyIleSerArGSerAlaThrIleAlaIleAlaIlyrIleMet 260
 DB 904 CTGATGCACGTGTTTACTGAGGACTCCCGCTCGGCACACATCCCTATCGCTATCATCATG 963
 QY 261 LysArgMetAaPSeArSerLeuAaPSeArGluAlaIlyrArGArPheValIysGluLysArProThr 280
 DB 964 AAGAGATGGACATGATCTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAACCTACT 1023
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnIleLeuAspIlyrGluLysIleLysAsn 300
 DB 1024 ATATCTCCAAACTTCAATTTCTGGGCAACCTCTGACATGAGAAAGATTAAGAAC 1083
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGluLysProAsn 320
 DB 1084 CAGACTGAGACATCAAGGCCCAAGAGCAAACTCAAGCTGTGCACCTGAGAGAGCCAAAT 1143
 QY 321 GluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProPro 340
 DB 1144 GAACGTGCTCCCTGCTGTCTCAGAGGGTGAAGAAAGCAAGAGCGCCCTCACTCCACCC 1203
 QY 341 CysAlaAaPSeArAlaHisSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 DB 1204 TGTGCGACTGTGCTACTCAAGAGCGACGACGAAAGGCCGTGATCCCGGACAGCTG 1263
 QY 361 ProSerValProSerValGlnProSerLeuLeuGluAaPSeArProLeuValGlnAlaLeu 380
 DB 1264 CCGAGCGTCCAGCGCTGAGCGCTGCTGTGTAAGAGACAGCCGCTGTGACAGGCGCTC 1323
 QY 381 SerGlyLeuHisLeuSerAlaAaPArgLeuGluAaPSeArAsnLysLeuLysArSerPhe 400
 DB 1324 AGTGGGCTGCACTGTCCGACAGCTGAGAAAGACAGCAATATGAGCTCAAGCGTTCCTTC 1383
 QY 401 SerLeuAspIleLysSerValSerIlyrSerAlaSerMetAlaLysLeuHisGlyLysPhe 420
 DB 1384 TCTCTGATATCAAAATCAGATTTATATTCAGCGAGATGCGACATCTTACATGACCTTC 1443
 QY 421 SerSerSerGluAspAlaLeuGluIlyrIlyrLysProSerThrThrLeuAspGlyThrAsn 440

Db 1444 TCCTCATCAGAAAGTCTTGGAAATACACAACTTCACACTCTGATGGAGCCAC 1503
QY 441 LysLeuCyGlnPheSerProValGlnGluSerGlnGlnThrProGlnThrSerPro 460
Db 1504 AGCATATGCAATCTCTCCCTCTTCAAGAACTATCGAGCACTCCCGAACCAGTCTT 1563
QY 461 AspLysGlnGlnValSerLeuProLysLeuGlnThrAlaArgProSerAspSerGln 480
Db 1564 GATTAAGGAGGAGGAGCAGATCCCAAGAGCTGGAGACCGCAGGCTTCAGACAGCCAG 1623
QY 481 SerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db 1624 AGCAGCAATTCATTCATTCGCTGACAAACAGCAGCAGTGGCAGCCGCAAGGCTCTTTA 1683
QY 501 SerProLeuHisArgSerGlySerValGlnAspAsnThrHisThrSerPheLeuPheGly 520
Db 1684 TCTCCACATCGATCGAAGTGGAGAGTGGAGACATTCACACACAGCTTCTTTTCGGC 1743
QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyThrHis 540
Db 1744 CTTTCACACAGCCAGCAGCACTTCACAGAACTGCTGGCTGGGCTTAAAGGCTGGCAC 1803
QY 541 SerAspLeuLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrProPheAla 560
Db 1804 TCGGATATCTGGCCCCCAGACCTTACCCCTTCCTGACACAGCTGGATATTTTGGC 1863
QY 561 ThrGlnSerSerHisPheThrSerAlaSerAlaIleThrGlyGlySerAlaSerThrSer 580
Db 1864 ACAGAGCTCTCACTCTTACTCTGCTCAGCCATCATCGAGGAGGAGTCCAGTTACTCT 1923
QY 581 AlaThrSerCysSerGlnLeuProThrCysGlyAspGlnValThrSerAlaArgArgArg 600
Db 1924 GCGTACAGCTGACGAGCCAGCTGCCACTTGGAGAGACCAAGTCTATCTGTGGCAGGCGG 1983
QY 601 GlnLysProSerSerAspArgAlaAspSerArgArgSerThrPheGlnGluSerProPheGln 620
Db 1984 CAGAGCGCAGTACAGACAGCTGACTCGCGCGAGCTGAGTGAAGAGAGCCCTTTGAA 2043
QY 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlnGlnSerIleMetSerGlnAsn 640
Db 2044 AAGCAGCTTTAAAGCGAGAGGCTGCCAAATGGAATTTGAGAGAGCATATATCGAGAAC 2103
QY 641 ArgSerArgGlnGlnLeuGlnGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
Db 2104 AGGTACCGGGAAGAGCTGGGGAAGTGGGAGTCACTTACGTTTTCGGGAGCATGGA 2163
QY 661 IleIleGlnValSer 665
Db 2164 ATCATTGAGTCTCC 2178

RESULT 14
AB051487
LOCUS AB051487 4790 bp mRNA linear PRI 07-FEB-2001
DEFINITION Homo sapiens mRNA for KIAA1700 protein, partial cds.
ACCESSION AB051487
VERSION AB051487.1 GI:12697944
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS [Nagase,T., Kikuno,R., Hatori,A., Kondo,Y., Okumura,K. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XIX. The complete sequences of 100 new cDNA clones from brain which
code for large proteins in vitro
JOURNAL DNA Res. 7 (6), 347-355 (2000)
MEDLINE 21082932
PUBMED 11214970
REFERENCE 2 (bases 1 to 4790)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,

Department of Human Gene Research, 1512-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna1@kazusa.or.jp)
URL: <http://www.kazusa.or.jp/huge>, Tel:81-438-52-3913,
Fax:81-438-52-3914)

FEATURES
source
location/Qualifiers
1..4790
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fj15353"
/note="vector:pb1uescriptII SK plus"
1..4790
/gene="KIAA1700"
<109..2181
/gene="KIAA1700"
/note="Start codon is not identified."
/codon_start=1
/product="KIAA1700 protein"
/protein_id="BA821791.1"
/db_xref="GI:12697945"
/translation="AFETHTITTSWQIEKGGRGLIVMAHEMIGTOIVTERVALT
ESGTERVLIDSRPFVENVNTSHLEAININGKMKRRLQODKVLITELIORSARKY
DIDSGKVVYDOSDQVSLSDPCLTVLGLKGRSPSHILAGFAERSRCPGL
CEGKSTLVPTCTISQPLPYANIGPRILIPNLYLQCDVINKELMQNGIGVPLVANS
TCPKPPIESPSPHRLRPVVDSPCEKILPMLDSVPFIRKASNGCVLHICLAGISRS
ATTALIVIKRMDSLDEAYRVRKRPITSPENFPLQQLDYERKIKINOTGASGPKS
KMLULHLEKNEPVPVVSQGSSTPLSPCADSATGAAGQRPVHPASVPSVQ
PSLLESPPLVQALSGHLISADRLSDNKLKSFSLDLSGVSASMAALSLGPFSSSD
ALEYKPSSTLDTGINKLQFSPVQELISQTPETSPDKENASIPKLTOTARPDSQSKR
LHSVRSSSGTRKORSLSPHRSQGVENVNHTSPFLRGISTCOHLITRAGLIGMS
DILAPOTSTPLTSSMYFATESSHFYSASALVYGSASISATISGQLPLCGQVSVSR
RQKSPDRARSRSWMBESPFEKQFRKSCQMFESJWNSRREELKGVGSQSSFSG
SMEILIEVS"

ORIGIN
Alignment Scores:
Pred. No.: 6,61e-204 Length: 4790
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
Gaps: 0
US-10-029-345A-109 (1-665) x AB051487 (1-4790)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuGln 20
Db 184 ATGGCCCATGAGATGATTTGAACTCAAAATGTTACTAAGAGTTGGGCTCTGCTGGA 243
QY 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnThrAsnThrSer 40
Db 244 AGTGGAAAGGAAAAAGCTGCTAATTAGCCGCGCATTTGTGAAATCAATATCACTCC 303
QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGlnGln 60
Db 304 CACATTTGGAAAGCATTAATATCACTGCTCAAGCTTATGAAGCGAAGTTGCAACAG 363
QY 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisValAspIleAsp 80
Db 364 GACAAAGGTTAATTAACAGAGCTCATCCAGCATTCAGCAACATAGGTTGACATGAT 423
QY 81 CysSerGlnLysValValValIleThrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 424 TGCAGTAGAAGGTGTGATGTTACGATCAAAAGCTCCCAAGTGGTCCCTCTCTTCA 483
QY 101 AspCysPheLeuThrValLeuLeuGlnLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db 484 GACGTTTCTCACTTACTCTCGGTTAACTGGAGAAAGCTTCAACTCTGTTCACCTG 543
QY 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuCyGlnGlyLysSer 140
Db 544 CTTGACAGGTGGTGGTTCGATGATTCCTCGTGTTCCTCGGCTCTGTGAAGAAATCC 603

QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 604 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACTTGTCCAAATTTGGGCCAAC 663
 QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGlyLeu 180
 DB 664 CGAATCTTCCCAATCTTTATCTTGGCTGCAGAGAAATGCTCTCAACAGAGAGCTGATG 723
 QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProlyProAspPhe 200
 DB 724 CACAGAAATGGAAATGGTATGTTAAATGACAGCAATACCTGTCCAAAGCCTGACTTT 783
 QY 201 ILProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIleLeu 220
 DB 784 ATCCCGAGCTCAATTTCTGCGTGTGCTGTGAATACAGCTTTGTGGAATAATTTTG 843
 QY 221 ProTyrLeuAspIleSerValAspPheIleGlnIleValAlaIleValAlaSerVal 240
 DB 844 CCGTGGTTGGCAAAATCAGTAGATTTCTATTGAGAAACAAAGCCTCAATGATGTT 903
 QY 241 LeuValHisCysIleLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 DB 904 CTATGTGACATGTTAGCTGGGAATCTCCGCTCCGCCAATCCCTATCGCTTACATCATG 963
 QY 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlnIleValAspProThr 280
 DB 964 AAGAGATGACATGTTTATGATGAACTTACAGATTTGTGAAAGAAAAGACCTACT 1023
 QY 281 ILSerProAsnPheAsnPheLeuGlyGlnLeuAspTyrGlnIleValIleValAsn 300
 DB 1024 ATATCTCCAAATCTTCAATTTTCTGGGCCAACTCCTGCACTATGAGAAAGATTAAAGAAC 1083
 QY 301 GlnThrGlyAlaSerGlyProIleSerIleLeuIleValLeuIleValIleValIleValProAsn 320
 DB 1084 CAGACTGAGACATCAGGCGCAAGCAAACTCAAGCTGTGCACTCGAGAAAGCCAAAT 1143
 QY 321 GlnProValProAlaValIleSerGlnIleGlyGlnIleSerGlnIleValProIleAspProPro 340
 DB 1144 GAACCTGTCTCTGTCTGACAGAGGGGAGACAGAAAGGCAAGAGAGCCCTCAGTCAAGCC 1203
 QY 341 CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 DB 1204 TGTGGGACTCTGCTACTCAGAGGAGCAGAGCAAAAGCCGTGACTCCGCAAGGTG 1263
 QY 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProIleValGlnAlaLeu 380
 DB 1264 CCCAGCCTGCCAGCGTGCAGCCGCTGTGAGAGCAGCCGCTGAGTACAGGCGCTC 1323
 QY 381 SerGlyLeuHisIleSerAlaAspArgLeuGlnAspSerAsnIleValIleValIleValIleVal 400
 DB 1324 AGTGGGCTGACCTGTCCGAGAGAGGCTGAAAGACAGCAATAGCTCAAGCGTCTTC 1383
 QY 401 SerLeuAspIleLeuSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 DB 1384 TCTCTGATATCAAAATAGTTTCAATATTCAGCAGCAGTGCAGATCTTACATGCTTC 1443
 QY 421 SerSerSerGlnAspAlaLeuGlnTyrTyrIleAspProSerThrIleAspGlyThrAsn 440
 DB 1444 TCTCTCATCAAGAAAGATGCTTGGAAATCTCAAACTTCCACTGTGATGGAGCAAC 1503
 QY 441 LysLeuCysGlnIlePheSerProValGlnGlnLeuSerGlnIleValProGlnIleValPro 460
 DB 1504 AAGCTATGCAAGATCTCCCTGTTCAAGAACTATCGAGAGAGCTCCCGAAACCAAGTCT 1563
 QY 461 AspIleGlnGlnIleAspIleProIleValLeuGlnIleThrAlaArgProSerAspSerGln 480
 DB 1564 GATTAAGAGAGAAAGCAGCATCTCCCAAGAGCTGAGAGCCCGAGGCTTACAGACAGCAG 1623
 QY 481 SerIleValIleHisIleSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 DB 1624 AGCAAGAGATTCATTCGTCTCAGAACAGAGAGCTGAGCAGCCGCGCAAGAGTCTCTTTTA 1683
 QY 501 SerProIleHisArgSerGlySerValGlnAspAsnIleValIleValIleValIleValIleVal 520

DB 1684 TCTCCACTCATCGAAGTGGAGCGTGGAGCAATATACACAGCACTTCTCCGCG 1743
 QY 521 LeuSerThrSerGlnGlnHisIleuThrIleValSerAlaGlyLeuGlyIleuIleValIleVal 540
 DB 1744 CTTTCCACAGCAGCAGAGCAGCTCAGAAAGTGTGCGGCTGGGCTTAAAGGCTGGCAC 1803
 QY 541 SerAspIleLeuAlaProGlnIlePheSerThrProSerLeuThrSerSerTyrPheAla 560
 DB 1804 TCGGATATCTGGCCCCCAGACCTTACCTTACCCTTCAACAGAGCTGGATTTTGCC 1863
 QY 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyIleValSerAlaSerTyrSer 580
 DB 1864 ACAGACTCTTCACTTACTTACTGTGCTGACGACATTAAGAGAGGAGTGCACATTA 1923
 QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValAlaGlyArg 600
 DB 1924 GCTTACAGCTGACGACGCTGCCCTTCCGAGACCAAGTCTATCTGTGCGCAGGCGG 1983
 QY 601 GlnIleProSerAspArgAlaAspSerArgArgSerTyrHisGlnIleValIleValIleVal 620
 DB 1984 CAGAACCAAGTACAGAGCTGACTGCGCGAGCTGGCATGAAAGAGAGCCCTTTGAA 2043
 QY 621 LysGlnPheIleValArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
 DB 2044 AAGCACTTTAAACGCAAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCAAGAAC 2103
 QY 641 ArgSerArgGlnGlnIleuGlyIleValGlySerGlnIleSerPheSerGlySerMetGln 660
 DB 2104 AGGTCAAGCGGAAGAGCTGGGGAAGAGTGGGACAGTCAAGCTTTTCGGGCAAGATGAA 2163
 QY 661 IleIleGlnValSer 665
 DB 2164 ATCATTTAGAGTCTCC 2178
 RESULT 15
 AX713989 2102 bp DNA linear PAT 15-Apr-2003
 LOCUS AX713989
 DEFINITION Sequence 673 from Patent EP1293569.
 ACCESSION AX713989
 VERSION AX713989.1 GI:29888917
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 1 Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuno, Y.
 TITLE Full-length cDNAs
 JOURNAL Patent: EP 1293569-A 673 19-MAR-2003;
 Helix Research Institute (JP) ; Research Association for Biotechnology (JP)
 FEATURES
 source location/Qualifiers
 1..2102
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,74e-204 Length: 2102
 Score: 3399.00 Matches: 662
 Percent Similarity: 99.70% Conservative: 1
 Best Local Similarity: 99.55% Mismatches: 2
 Query Match: 99.44% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-345A-109 (1-665) x AX713989 (1-2102)
 QY 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20

Db	56	ATGGCCCAAGAGATGATTGGAACTCAAAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGGAA	115
OY	21	SergLYthGluYsValLeuLeuIleAspSerArgProPheValGluTyraSer	40
Db	116	ACTGGAAACGGAAAGAGTGGCTCAATTGTGATACGGCGCAATTTGTGGAAATCAATACATCC	175
OY	41	HisIleLeuGlnAlaIleAsnIleAsnCyserIysLeuMetLysArgArgLeuGlnG	60
Db	176	CACATTTTGGAGGCCATTAAATCACTCTCCACCTTATGAGAGGAAAGTTGGCAACG	235
OY	61	AspLYsValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp	80
Db	226	GACAAAGTGTAAATTAACAAGGTCATCCAGCATTCAGCAAAACATTAAGTTACATTTGAT	295
OY	81	CysSerGlnLysValValValTyraAspGlnSerSerGlnAspValAlaSerLeuSerSer	100
Db	236	TGCACATCAGAAAGGTTGAGTTTACATTCACATCAAACTCCCAAGATGTGGCTCTCTCTTCA	355
OY	101	AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu	120
Db	356	GACGTGTTTCTACAGTACTTCTGGGTAACAGAGAAAGCTTCAACTCTGTACACTG	415
OY	121	LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGlnGlyLysSer	140
Db	416	CTTGCAGGGGGGTTTGCTGAGATTCTCTGTTTTCCTCGGGCTGTGAAAGAAATCC	475
OY	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	476	ACTCTAGTCCCTACCTGCACTTCTCAGGCTTGTACCTGTTGGCCAACTTGGGCCAAC	535
OY	161	ArgIleLeuProAsnLeuTyrlengIlyCysGlnArgAspValIleAsnIlySgIleuIle	180
Db	536	CTAAATCTTCCCAACTTATCTTGGCTGCGACGACGAGATGTCCTCAACAGAGAGCTGAG	595
OY	181	GlnGlnAsnGlyIleGlyTyrlValLeuAsnAlaSerTyrlThrCysProLysProAspPhe	200
Db	556	CAGCAGAATGGGATTTGATGTGTTAAATGACACAAATACCTGTGCCAAAGGCTCACTTT	655
OY	201	IleProGlnSerHisPheLeuArgValProValaAsnAspSerPheCysGlnLysIleLeu	220
Db	656	ATCCCGAGTCTCATTTCTCGCTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG	715
OY	221	ProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal	240
Db	716	CCGTGGTTGGCAAACTCAGTAAATTTCATTGGAAGCAAAAGCCTCCAAATGAGTGT	775
OY	241	LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlLeuMet	260
Db	776	CTAGTGCACCTGTTTAGCTGGATCTCCCGCTCCGCACCATGTCTATGGCTCAATCATG	835
OY	261	LysArgMetAspMetSerLeuAspGlnAlaTyraArgPheValLysGlnLysArgProThr	280
Db	836	AAGAGGATGACATGTCTTTGATGAACCTTACATTTGTGAAGAAAGAAAGCACTTCT	895
OY	281	IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrlArgLysIleLysAsn	300
Db	896	ATATCTCCAAACTTCATTTTCTGGGCAACTCTCGGACTGAGCAAGAAAGATTTAAGAAC	955
OY	301	GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlnLysProAsn	320
Db	956	CAGACTGAGCATCAGGGCCAAAGACCAAACTCAAGCTGTCTGCACCTTGAGAGGCAAAAT	1015
OY	321	GluProValProAlaValSerGlnGlyGlyGlnLysSerGluThrProLeuSerProPro	340
Db	1016	GAACCTGTCCCTGTCTCAGAGGGTGGACAGAAACCGAGACCCCTTCATGTCACACC	1075
OY	341	CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal	360
Db	1076	TGTGCGCACTCTGCTACTCAGAGGACAGACAGCAAAAGCCCGTGTCAATCCGCCAGCGTG	1135
OY	361	ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu	380

[illegible]

Search completed: June 22, 2004, 02:43:11
Job time : 7843.33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: June 21, 2004, 21:18:06 ; Search time 819.731 Seconds

(without alignments)
3716.370 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 3418
Sequence: 1 MAHEMGTGIVTVERLVALE.....LKGVSQSSFSGSMETIEVS 665

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10029345/runcat_21062004_16416_29071/app_query.fasta_1.1294
-DB=Published Applications NA -QFMT=fastbp -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cst -LIST=45 -DOCAIGN=200 -THR.SCORE=pcr -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPEXT=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029345 -CGN_1_1_333 @runcat_21062004_16416_29071
-NCPV=6 -ICPV=3 -NO MMAP -LARGEDUERY -NEG SCORES=0 -WAIT -DSBBLCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPEXT=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
------------	-------------	--------	----	-------------

1	3406	99.6	1998	6	US-09-816-494-3	Sequence 3, Appl1
2	3406	99.6	1998	9	US-10-377-072-27	Sequence 27, Appl1
3	3406	99.6	2732	13	US-10-168-506-2	Sequence 2, Appl1
4	3406	99.6	3059	17	US-10-257-028-1	Sequence 1, Appl1
5	3406	99.6	3496	9	US-09-964-277-1	Sequence 1, Appl1
6	3406	99.6	3544	9	US-09-816-494-1	Sequence 1, Appl1
7	3406	99.6	3544	16	US-10-377-072-25	Sequence 25, Appl1
8	3406	99.6	3625	13	US-10-425-114-26234	Sequence 26234, A
9	3406	99.6	3766	13	US-10-343-357-17	Sequence 17, Appl1
10	3406	99.6	4790	17	US-10-648-559-115	Sequence 115, Appl1
11	3399	99.4	2102	16	US-10-094-749-673	Sequence 673, Appl1
12	3398	99.4	2866	13	US-10-296-115-520	Sequence 520, Appl1
13	3379.5	98.9	2071	13	US-10-072-012-257	Sequence 257, Appl1
14	3358.5	98.3	2200	13	US-10-072-012-255	Sequence 255, Appl1
15	3048.5	89.2	3332	9	US-09-964-277-20	Sequence 20, Appl1
16	1643.5	48.1	1916	16	US-10-108-260A-2429	Sequence 2429, Appl1
17	1302	38.1	2453	14	US-10-005-858-1	Sequence 1, Appl1
18	1262	36.9	2476	13	US-10-220-120-28	Sequence 28, Appl1
19	1075.5	31.5	2039	13	US-10-072-012-265	Sequence 265, Appl1
20	480	14.0	2109	16	US-10-305-720-1135	Sequence 1135, Appl1
21	480	14.0	2109	17	US-10-641-643-946	Sequence 946, Appl1
22	480	14.0	2390	13	US-10-342-887-698	Sequence 698, Appl1
23	480	14.0	2390	13	US-10-172-118-698	Sequence 698, Appl1
24	480	14.0	2390	15	US-10-175-523-183	Sequence 183, Appl1
25	480	14.0	2390	16	US-10-199-221-4	Sequence 4, Appl1
26	480	14.0	2380	17	US-10-717-597-4	Sequence 4, Appl1
27	480	14.0	2792	10	US-09-971-392-69	Sequence 69, Appl1
28	479	14.0	2794	14	US-10-044-090-330	Sequence 330, Appl1
29	477	14.0	2649	16	US-10-133-937-48	Sequence 48, Appl1
30	477	14.0	2649	16	US-10-159-563-48	Sequence 48, Appl1
31	471	13.8	2650	15	US-10-184-832-1	Sequence 1, Appl1
32	469	13.7	1337	9	US-09-964-899-46	Sequence 46, Appl1
33	469	13.7	1449	12	US-10-052-488-192	Sequence 192, Appl1
34	469	13.7	1830	15	US-10-346-356-1	Sequence 1, Appl1
35	469	13.7	1909	12	US-10-052-488-191	Sequence 191, Appl1
36	469	13.7	1909	13	US-10-058-270A-129	Sequence 129, Appl1
37	468	13.7	1452	12	US-10-052-488-189	Sequence 189, Appl1
38	468	13.7	1977	12	US-10-052-488-188	Sequence 188, Appl1
39	466.5	13.6	1143	15	US-10-184-832-3	Sequence 3, Appl1
40	466.5	13.6	2104	12	US-10-152-319A-5041	Sequence 2041, Appl1
41	466.5	13.6	2104	16	US-10-388-934-654	Sequence 654, Appl1
42	466.5	13.6	2104	16	US-10-191-803-270	Sequence 270, Appl1
43	461	13.5	3286	10	US-09-971-392-9	Sequence 9, Appl1
44	453	13.3	1238	9	US-09-736-457-803	Sequence 803, Appl1
45	453	13.3	1238	9	US-09-902-941-803	Sequence 803, Appl1

ALIGNMENTS

RESULT 1
US-09-816-494-3
Sequence 3, Application US/09816494
Patent No. US20020034807A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
FILE REFERENCE: PHOSPHATASE MOLECULES AND USES THEREFOR
CURRENT APPLICATION NUMBER: US/09/816,494
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,858
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1998
TYPE: DNA
ORGANISM: Homo sapiens
US-09-816-494-3

Alignment Scores: 0 Length: 1998
Pred. No.: 0

Sequence 27, Application US/10377072
Publication No. US20040009501A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals Inc.
APPLICANT: Curtis, Rory A.J.
APPLICANT: Logan, Thomas Joseph
APPLICANT: Gluckmann, Maria A.
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Miyoung
APPLICANT: Teal, Fong-Ying
TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: MP03-0180NMIM
CURRENT APPLICATION NUMBER: US/10/377,072
PRIORITY FILING DATE: 2003-02-27
PRIORITY FILING DATE: 2001-06-29 US 09/895,860
PRIORITY FILING DATE: 2001-06-29 US 60/215,370
PRIORITY FILING DATE: 2000-06-29
PRIORITY FILING DATE: 2000-06-29 US 09/723,806
PRIORITY FILING DATE: 2000-11-28
PRIORITY FILING DATE: 2000-03-07 US 60/187,455
PRIORITY FILING DATE: 2000-03-07
PRIORITY FILING DATE: 2001-04-25
PRIORITY FILING DATE: 2001-04-25 US 60/199,801
PRIORITY FILING DATE: 2000-04-26
PRIORITY FILING DATE: 2001-05-21 US 09/861,801
PRIORITY FILING DATE: 2001-05-21
PRIORITY FILING DATE: 2000-05-19 US 60/205,508
PRIORITY FILING DATE: 2000-05-19
PRIORITY FILING DATE: 2001-03-23 US 09/816,494
PRIORITY FILING DATE: 2001-03-23
PRIORITY FILING DATE: 2001-03-22 US 09/815,419
PRIORITY FILING DATE: 2001-03-22
Remainder Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 1998
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1998)
US-10-377-072-27
Alignment Scores:
Pred. No.: 0 Length: 1998
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
Gaps: 0
US-10-029-345a-109 (1-665) x US-10-377-072-27 (1-1998)
QY 1 MetAlaHisGluWettIegIyThrGlnIleValThrGluuArgLeuValAlaIleuLeuGlu 20
DB 1 ATGGCCCATAGATGATTGGAACCAATTGTTACTGAGAGTTGGTGGCTGCTGGAA 60
QY 21 SerGIyThrGluuValIleuLeuIleAspSerArgProPheValGluTyraAntThrSer 40
DB 61 AGTGGAGCGAAAGAGCTGCTGATGATGACCGGCCATTTGGAGATACATACATCC 120
QY 41 HisIleuGluGluAlaIleAsnIleAsnCyseSerIyLeuWetiIyValArgLeuGluGln 60
DB 121 CACATTTTGGAGGCAATTAATCACTGCCCAAGCTTATGAAGCGAAGTTGCAACAG 180
QY 61 AspIyValIleuIleThrGluLeuIleGlnHisSerIyAlaIyHisIyValAspIleAsp 80

DB 181 GACAAAGTGTATTAATCAGAGCTCATCCAGCATTCAGCGAAACATAGTTGACATTGAT 240
QY 81 CyseSerGluIyValIleValIyAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 241 TGCAGTCAGAGAGTTGATGATTCAGATCAAGCTCCCAAGATTTGCTCTCTCTTCA 300
QY 101 AspCyPheLeuThrValIleuGluIyIyLeuGluIySerPheAsnSerValHisIleu 120
DB 301 GACTGTTTCTCACTACTCTGAGGTAACTGGAAGAGGCTTCACTGTTCACTG 360
QY 121 LeuAlaGlyIyPheAlaGluPheSerArgCyPheProGlyIyLeuCyGluGlyIySer 140
DB 361 CTTCGAGGAGGAGTTCTGATGTTCTCTGTTTCCCTGCGCTCTGTAGAGAAATCC 420
QY 141 ThrLeuValProThrCyHisSerGlnProCyLeuProValAlaAsnIleGlyProThr 160
DB 421 ACTCTAGTCCCTACCTGACATTTCTCAGCCCTTGCTTCTGCTTCCCAACATGGGCCAACCC 480
QY 161 ArgIleLeuProAsnLeuTyrlleuGlyCyGlnArgAspValIleuAsnIyGluIle 180
DB 481 CGAATCTTCCCAATCTTATCTTGCTGCGCAGCGAGATGCTCTCAACAGAGCTGATG 540
QY 181 GlnGlnAsnGlyIleGlyTyrlValIleAsnAlaSerTyrlThrCyAspProIyAspPhe 200
DB 541 CAGCAGAAATGGATTTGTTATGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 600
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCyGluIyIleLeu 220
DB 601 ATCCCGAGCTCATTTCTGCGGTGCTGTGAAGCAGCTTTGTGTGAGAAATTTTG 660
QY 221 ProThrLeuAspIySerValAspPheIleGluIyAlaIyValAsnGlyIyCyVal 240
DB 661 CCGTGTGTGACAAATCAGTAGTTTCATTGAAAGCAAAACCTTCAATGATGTGTT 720
QY 241 LeuValHisCyPheLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlIleMet 260
DB 721 CTGATGCACTGTTTACTGATGATGATCTCCGCTCCGCAACATGCTTACCTCATCATG 780
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrlArgPheValIyGluIyValArgProThr 280
DB 781 AAGAGATGAGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAACCACTACT 840
QY 281 IleSerProAsnPheAsnIleuGluIleLeuAspTyrlGluIyIyValIleAsn 300
DB 841 ATATCTCCAAACTTAATTTCTGGGCAACTCTGACCTATGAGAGAAATTAAGAAC 900
QY 301 GlnThrGlyAlaSerGlyProIySerIyLeuIyLeuIleuHisIleuGluIyProAsn 320
DB 901 CAGACTGAGACATCAGGCGCAAGAGCAAACTCACTGCTGACCTGAGAGAGCCAAAT 960
QY 321 GluProValProAlaValSerGluGlyIyGluIySerGluThrProLeuSerProPro 340
DB 961 GAACTGTCTCTGCTCTCTCAGAGGAGTGAAGAAAGCGAGCGCCCTCAGTCCACCC 1020
QY 341 CyValAspSerAlaHisSerGluAlaIleGlyIleArgProValHisProAlaSerVal 360
DB 1021 TGTGCGACTGCTGCTACCTCAGAGGAGCAGCAAGAGCCCTGTGATCCCGCAGCGTG 1080
QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
DB 1081 CCGAGCTGCCCAAGCTGAGCGCTGCTGTTTGAAGAGCAGCCGCTGTGACAGCGCTC 1140
QY 381 SerGlyLeuHisIleuSerAlaAspArgLeuGluAspSerAsnIyLeuIyValArgSerPhe 400
DB 1141 AGTGGCTGACCTGCTCCGAGCAGGCTGGAAGACAGCAATAGAGCTTCAAGCTTCTTC 1200
QY 401 SerLeuAspIleIySerValSerTyrlSerAlaSerMetAlaIleSerIleuHisIleGlyPhe 420
DB 1201 TCTCTGATATATCAAACTATTCATATTCAGCCAGATGCGACATCTTACATGCTTTC 1260
QY 421 SerSerSerGluAspAlaIleuGluTyrlTyrlProSerThrThrLeuAspGlyTyrlAsn 440


```

Db      1261 TCCTCATCGAAGATGCTTGGAAATACAAACCTTCCACTACTGTGATGGAGCAAC 1320
Qy      441 LysLeuCySGlnPheSerProValGlnGluLeuSerGlnGlnThrProGlnThrSerPro 460
Db      1321 AAGCTATGCGAGTCTCCCTCTGTCAGGAAGCTACGAGAGCACTCCCAAAACATCTCT 1380
Qy      461 AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
Db      1381 GATAAAGAGAAAGCAAGCATCTCCCAAGAAAGCTCAGACCGCCAGGCTTCAGACAGCCAG 1440
Qy      481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db      1441 AGCAGAGATTGCAATTCGTCAAGAACACAGACAGAGTGGACCGCCCAAGAGTCTCTTTTA 1500
Qy      501 SerProLeuHisArgSerGlySerValGluAspArgTrpHisThrSerPheLeuPheGly 520
Db      1501 TCTTCACCTGCATCGAAGTGGAGCGGAGAGCAATTACACACAGACTCTCTTTTGGCC 1560
Qy      521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTrpHis 540
Db      1561 CTTTCCACAGCCAGCAGACCACTCAGAAAGTCTGCTGGCTTGAAGGCGCTGGCAC 1620
Qy      541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpYrPheAla 560
Db      1621 TCGAATATCTTGCCCCCAGACCTTACCCCTTCCCTGACACAGCACTGTATTGTC 1680
Qy      561 ThrGlnSerSerHisPheThrSerAlaSerAlaIleTrpGlyGlySerAlaSerTrpSer 580
Db      1681 ACGAGTCTCTACACTTCTACTCTGCTCAGCCATCTACGAGGCAAGTGCACATTACTCT 1740
Qy      581 AlaTrpSerCySerGlnLeuProThrCyGlyAspGlnValTrpSerValArgArgArg 600
Db      1741 GCTTACAGCTCAGCGCAAGCTCCCACTTGGGAGACCAAGTCTATTCTGTGGCAGGCGG 1800
Qy      601 GlnLysProSerAspAlaGlnAlaAspSerArgArgSerTrpHisGlnGlnSerProPheGlu 620
Db      1801 CAGAGCGCAAGTGCACAGCTGACTCGCGCGAGGCTGGCAGTGAAGAGAGCCCCCTTTAA 1860
Qy      621 LysGlnPheLysArgArgSerCySGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
Db      1861 AAGCAGTTTAAACGAGAAAGCTCCAAATGGAATTTGGAGAGCATCATGTCAGAGAAC 1920
Qy      641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db      1921 AGGTCAAGGGAAGAGCTGGGGAAAGTGGCACTGCTGAGCTTTTGGGGCAGCATGAAA 1980
Qy      661 IleIleGlnValSer 665
Db      1981 ATCATTCAGGTCTCC 1995

```

RESULT 3

```

US-10-168-506-2
; Sequence 2, Application US/10168506
; Publication No. US20040053229A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: SUNDARSANAM, SUCHA
; APPLICANT: HILL, RON
; APPLICANT: PLANAGAN, PETER
; TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
; FILE REFERENCE: 038602/1351
; CURRENT FILING DATE: US/10/168, 506
; PRIOR APPLICATION NUMBER: 2002-06-21
; NUMBER OF SEQ ID NOS: 12-21
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2732
; TYPE: DNA

```

ORGANISM: Homo sapiens
US-10-168-506-2

Alignment Scores:

	Score:	Length:
Pred. No.:	0	2732
Percent Similarity:	3406.00	Matches: 663
Best Local Similarity:	99.85%	Conservative: 1
Query Match:	99.70%	Mismatches: 1
DB:	99.65%	Indels: 0
	13	Gaps: 0

US-10-029-345a-109 (1-665) x US-10-168-506-2 (1-2732)

```

Qy      1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db      538 ATGGCCCATGAGATGATTGGAACTCANATTTGTTACTGAGAGGTTGGTGGCTGCGGAA 597
Qy      21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTrpAsnThrSer 40
Db      598 AGTGAACGGAAAAAGCTGCTAATTCATAGCCGGCCATTGTGGATACATACATCC 657
Qy      41 HisIleLeuGlnAlaIleAsnIleAsnGlySerLysMetLysArgArgLeuGln 60
Db      658 CACATTTTGGAGCCATTATATCATCTGCTCAGAGCTTATGAAACGAAAGGTTGCAACAG 717
Qy      61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db      718 GACAAAGTTAATTAACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 777
Qy      81 CySerGlnLysValValValTrpAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db      778 TGCAGTCAGAAAGTTGATTAGCATTAAGTCCCAAGATGTGGCTCTCTCTCTTCA 837
Qy      101 AspCyPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db      838 GACTGTTTCTCAGTACTTCTGTGGTAAACTGGAGAAAGCTTCAACTCTGTTCACCTG 897
Qy      121 LeuAlaGlyGlyPheAlaGluPheSerArgCyPheProGlyLeuCyGlnGlyLysSer 140
Db      898 CTTCGAGGTGGGTTTGCTGAGTTCTCTGCTGTTTCCCTGGCTCTGTGAAGAAATCC 957
Qy      141 ThrLeuValProThrCySIIeSerGlnProCyLeuProValAlaAsnIleGlyProThr 160
Db      958 ACTCTAGTCCCTACTCTCATTTCTCAGCCCTGCTTACTGTTGCCAACATTTGGCCAACC 1017
Qy      161 ArgIleLeuProAsnLeuTrpLeuGlyCySGlnArgAspValLeuAsnLysGlnLeuIle 180
Db      1018 CGAATTCCTCCCAATCTTTATCTTGGCTCCAGCGAGAGTCTCTCAACAGAGAGCTGATG 1077
Qy      181 GlnGlnAsnGlyIleGlyTrpValLeuAsnAlaSerTrpThrCySProLysProAspPhe 200
Db      1078 CAGCAGATGGAGATGATGATGATGTTAAATGCCACAAATACCTGTCCAAAGCTGACCTTT 1137
Qy      201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCyGlnLysIleLeu 220
Db      1138 ATCCCGAGTCAATTCCTGCTGCTGCTGTGATGACACCTTTTGTGAAAAATTTTG 1197
Qy      221 ProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCyVal 240
Db      1198 CCGTGGTTGACAAATTCAGTAAATTCATTTCAATGAGAAACAAAGCTCCATGATGTGTT 1257
Qy      241 LeuValHisCySLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTrpIleMet 260
Db      1258 CTAGTGACACTGTTTAGTGGATCTCCCGCTCCGCCACCATTCCTATCCCTACATCAG 1317
Qy      261 LysArgMetAspMetSerLeuAspGlnAlaTrpArgPheValLysGlnLysArgProThr 280
Db      1318 AAGAGATGAGCATGCTTTAGATGAAGCTTACAGATTGTGTAAGAAAAGAAAGCACTACT 1377
Qy      281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTrpGlnLysLysIleLysAsn 300
Db      1378 ATATCTCCAAACTCATTTTTCGGGCACATCTCTGACATGAGAAAGATTAAAGAAC 1437

```

QY	301	GIlnthrgIyAlaSerGIyProIySeriTySleuIySleuIhIleuGIuIySProIaSn	320
Db	1438	CAGACTGAGATACAGGGCCAAAGGACCAACTCAAGCTGTGCACCTGGAGAAAGCCAAAT	1497
QY	321	GIuProValProAlaValSerGIuGIyGIyGIuIySerGIuIhProIeuSerProPro	340
Db	1498	GAACTGTCTCTGCTGTTCAGAGGGGTGACAGAAAGCGAGACGCCCTTCAAGTCCACC	1557
QY	341	CyAlaIaAPSerAlaThrSerGIuAlaAlaGIyGlnaIyGlnaIySProValIhIleProAlaSerVal	360
Db	1558	TGTGCCGACTGTGCTACTCAGAGGACAGACAGACAAAGGCCGTGATCCGCCAGCGGTG	1617
QY	361	ProSerValProSerValGlnProSerIleuGIuIaAPSerProIeuValGlnAlaIeu	380
Db	1618	CCGAGCGGCCACAGGTGCACCGCTGCTTGAAGGACACCGCTGTGTACAGGCGCTC	1677
QY	381	SerGIyIleuIhIleuSerAlaAPaArgIeuGIuIaAPSerAluIyIleuIyAlaIySArgSerPhe	400
Db	1678	AGTGGGCTGACCTGTCCGACAGACGGCTGGAGAGACAAATAGCTTAAGCTTCTTTC	1737
QY	401	SerIleuAPhIleIySArgSerValSerTySAlaSerMetAlaAlaSerIleuIhIleGIyPhe	420
Db	1738	TCTGTGGATTCAAATCATCGTTTCANATTCAACCCAGACAGAGAGATCTTTACATGGCTTC	1797
QY	421	SerSerSerGIuAPAlaIleuGIuTyTyIySProSerThrThrIleuAPGIyThraSn	440
Db	1798	TCCCTCATCAGAAAGATGCTTGTGAATACTCAAAACCTTCACCTACTCTGTGAAGGACCAAC	1857
QY	441	IyIleuCyAGIhIleSeriSeriProValGlnIuIleuSerGIuGIuIhProGIuIhThSerPro	460
Db	1858	AAGCTATGCCAGTTCTCCCTGTTTCAGGAACATTCGAGGACAGACTCCGAAACCAAGTCT	1917
QY	461	APhIySGIuGIuAlaSerIleProIyIyIleuGIuIhAlaArgProSerAPSerGIu	480
Db	1918	GATTAAGAGAGAAAGCCAGCATCCCAAGAAAGCTGCAGCATGCCAGGCTTTCAGACAGCCAG	1977
QY	481	SerIyAlaIleuIhIleSerValaIArgThSerSerSerGIyThraIaGlnaIySeriIleu	500
Db	1978	AGCAAGCATTTGCATTCGGTCAGAAACCAAGCAGCTGGACACCGCCAGAGGTCTCCCTTTTA	2037
QY	501	SerProIeuIhIaIySArgSerGIySerValGIuIaAPaPhIyTyIhIhThSerPheIleuPheGIy	520
Db	2038	TCTCCACTGCATCCAGATGGAGAGCTGGAGACAAATTACACACACGTTCTTCATTTCCGGC	2097
QY	521	IeuSerThrSerGIuGlnIhIleuThrIySArgSerAlaGIyIleuGIyIleuIyGIyThraIhI	540
Db	2098	CTTTCACACACCCAGACGACTCTACAGAAAGTCTGCTGGCCTTGAGGCTTAAAGGCTGGAC	2157
QY	541	SerAPhIleuAlaProGIuIhThSerThrProSerIleuIhIleSerSerTyTyPheAla	560
Db	2158	TGGATATCTTGAGCCGCCACAGCCCTTACCCCTTCGACACAGACGCTGGATTTTGC	2217
QY	561	ThrGIuSerSerThraPheTySArgSerAlaSerAlaIleTyGIyGIySerAlaSerTySer	580
Db	2218	ACAGAGTCTCTACACTTCTACTCTGCTCCACAGCACTTACGAGGAGGAGCCAGTTAACTCT	2277
QY	581	AlaTySerCySArgSerGlnIeuProThrCySArgIyAPGIuAlaTySerValaIArgArgArg	600
Db	2278	GCTTACAGCTCAGACGACGACTGCCACTTGCAGAGACCAAGTCTATTCTGTGCGAGGCGG	2337
QY	601	GlnIySProSerAPaArgAlaAPSerAlaIySArgSerTyPheIhIleGIuIeuSerProPheGIu	620
Db	2338	CAGAAAGCCAAAGTACAGAGCTGACTCCGGCGAGAGCTGGCATTAAGAGAGAGCCCTTTGAA	2397
QY	621	IyIleGIuIhIleIyAlaIySArgSerCyAGIleuIleGIuIhIleGIyGIuSerIleuSerGIuIaSn	640
Db	2398	AAGCAGTTTAAACCGAGAAAGCTGCCAAATGGAAATTGGAGAGAGCATCATGTCCAGAGAAC	2457
QY	641	ArgSerArgGIuGIuIeuGIyIyIyValaGIySerGlnSerSerPheSerGIySerMetGIu	660
Db	2458	AGGTACCGGAGAAAGCTGGGAAAGGTGGGAGTCAAGTCTTAAGCTTTTCGGGACGATGGAA	2517
QY	661	IleIleGIuIaIleSer	665

DB	2518	ATCATTGAGCTCC	2532
	RESULT 4		
	US-10-257-026-1		
	; Sequence 1, Application US/10257026		
	; Publication No. US20040086859A1		
	; GENERAL INFORMATION:		
	; APPLICANT: Merck Patent GmbH		
	; TITLE OF INVENTION: New dual specificity phosphatase		
	; FILE REFERENCE: DUSP10K0MS		
	; CURRENT APPLICATION NUMBER: US/10/257, 026		
	; CURRENT FILING DATE: 2003-11-07		
	; NUMBER OF SEQ ID NOS: 3		
	; SOFTWARE: PatentIn Ver. 2.1		
	; SEQ ID NO 1		
	; LENGTH: 3059		
	; TYPE: DNA		
	; ORGANISM: Homo sapiens		
	; FEATURE:		
	; NAME/KEY: CDS		
	; LOCATION: (127)..(2121)		
	US-10-257-026-1		
	Alignment Scores:		
	Pred. No.: 0	Length: 3059	
	Score: 3406.00	Matches: 663	
	Percent Similarity: 99.85%	Conservative: 1	
	Best Local Similarity: 99.70%	Mismatches: 1	
	Query Match: 99.65%	Indels: 0	
	DB: 17	Gaps: 0	
	US-10-029-345A-109 (1-665) x US-10-257-026-1 (1-3059)		
QY	1	MetAlaHISgluMetIIEGLYThrGlnIleValThrGluArgGluValAlaLeuLeuGlu	20
DB	127	ATGGCCCATGACATCATTTGGAACTCAAAATTTGTTACTGAAAGTTGGTGGCTGTGCGAA	1868
QY	21	SerGIYThrgIuLyVaIleuLeuIleAspSerArGProPheValGluTyraenThSer	40
DB	187	AGTGAAACGAAAAAGGTGTGCTTAATTGATTAGCCGGCCATTGTGTGAATACATATCATCC	2468
QY	41	HisIleLeuGluAlaIleAsnIleAsnCySerLySleuMetLyArGArgLeuGlnGln	60
DB	247	CACATTTTGGAAAGCCATTATATCACTGCTCCAGCTTATGAAAGGAAGTTGCACACG	3068
QY	61	AspLySValIleuIleThrGlnuLeuIleGlnHisSerAlaIyHisLySValIleAsp	80
DB	307	GACAAAGGTATTAATACAGAGCTCATCCAGCATTCAGCAAAATAAAGTTGACATTTGAT	3668
QY	81	CysSerGlnLySValIleValTyraAspGlnSerSerGlnAspValAlaSerLeuSerSer	1008
DB	367	TGCAATCAGAAAGTTGTACTTATTCATCAAGCTCCCAAGATGTTGCCCTCTCTCTTCA	4268
QY	101	AspCyPheLeuThrValIleuLeuGlyLySleuGluLySerPheAsnSerValIleu	1208
DB	427	GACTGTITTTCTCACTGACTTCTGGGTAACGAGAAAGACTTCAACTCTGTTCACCTG	4868
QY	121	LeuAlaGlyGlyPheAlaGluPheSerArGyPheProGlyLeuCyGluGlyLySer	1408
DB	487	CTTGAGAGGGGTGTGAGTTCCTCGTGTITTTCCCTGGCCCTGTGTGAAGAAATCC	5468
QY	141	ThrLeuValProThrCysIleSerGlnProCySleuProValAlaAsnIleGlyProThr	1608
DB	547	ACTCTAGTCCCTCACTGATTTCTCAGCCTTCTTAACGTGTGCCAAATTTGGGCCAAC	6068
QY	161	ArgIleLeuProAsnLeuTyrluGlyCySgluArgAspValIleuAsnLySgluLeuIle	1808
DB	607	CGAATTTCTTCCCAATCTTATCTTGGCTGCGACGCGAGATGTCCTCAACAAGAGAGCTGAT	6668
QY	181	GlnGlnAsnGlyIleGlyTyrrValIleuAsnAlaSerTyrrThCySProLySProAspPhe	2008
DB	667	CAGCGAATGGATGGATTGATGTGTTAAATGCGAGCAATACCTGTCCAAAGCTTGACTTT	7268

```

QY 201 ILeProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluYsileu 220
DB 727 ATCCCGAGATCATTTCTCGTGCTGTGATGACAGCTTTGTGGAATAATTTTG 786
QY 221 ProTribLeuAspLysSerValAspPheIleGluYsAlaYsAlaSerAnGlyCysVal 240
DB 787 CCGTGGTGGCAAAATAGTATGATTCATTGAGAAAGCAAAAGCTCCATGAGATGTGT 846
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
DB 847 CTAGTGCATCTTTAGTGGATCTCCCGCTCCGACCATGCTATGAGAAAGATTAGAAC 906
QY 261 LysArgMetAspMetSerLeuAspGluAlaIleArgPheValIleGluYsAlaArgProThr 280
DB 907 AAGGAGATGAGACATGCTTTAGATGAGAGCTTACAGTTTGTGAAGAAAAAGACCTACT 966
QY 281 ILeSerProAsnPheAsnPheLeuGlyIleLeuAspIleGluYsAlaIleYsAsn 300
DB 967 ATATCTCCAACTTCAATTTTCTGGCCAACTCTGGACTATGAGAAAGATTAGAAC 1026
QY 301 GluThrGlyAlaSerGlyProLysSerIleLeuYsLeuHisIleGluYsProAsn 320
DB 1027 CAGACTGGAGACATCGAGGCGCAAAAGAGCAAACTCAAGGCTGCTGACCTGGAGAAAGCCAAAT 1086
QY 321 GluProValProAlaValSerGluGlyIleGluYsSerGluThrProLeuSerProPro 340
DB 1087 GAACCTGTCCTGCTGCTCAGAGGGTGGACAGAAAGCAGACGCCCCCTCAGTCCACCC 1146
QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyIleGluYsProValHisProAlaSerVal 360
DB 1147 TGTGCGACTCTGCTACTCTCAGAGCGACAGCAAAAGCCGCTGCTACAGCGCTCC 1206
QY 361 ProSerValProSerValGluProSerLeuLeuGluAspSerProLeuValGluAlaLeu 380
DB 1207 CCCAGCGGCCCAACGTCGACCGCTGCTTGAAGAGACGCCCCGCTGCTACAGCGCTCC 1266
QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIleLeuYsArgSerPhe 400
DB 1267 AGTGGCTGCACCTGCTCGCAGACAGCGCTGGAAGCAGCATATACCTCAACGCTTCTTC 1326
QY 401 SerLeuAspIleYsSerValSerIleSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
DB 1327 TCTCTGATATCAAAATCAGTTTCAATTCAGCCACAGATGCGCATCTTACATGCTTC 1386
QY 421 SerSerSerGluAspAlaLeuGluIleYsIleYsProSerThrThrLeuAspGlyThrAsn 440
DB 1387 TCTCTATCGAAGATGCTTTGGAATATCAAAACCTTCACTGCTGAGTGGACCAAC 1446
QY 441 LysLeuCysGluPheSerProValGluGluLeuSerGluGluThrProGluThrSerPro 460
DB 1447 AAGCTATGCGCAAGTCTCCCTGTTCAAGCACTATCGAGAGCATCTCCGAAACCAAGTCTC 1506
QY 461 AspLysGluGluAlaSerIleProLysIleLeuGluThrAlaArgProSerAspSerGln 480
DB 1507 GATTAAGAGAGAACGAGCATCCCAAGAGCTGAGACCGCCAGGCTTCAAGACCAAG 1566
QY 481 SerIleValGluHisSerValArgThrSerSerSerGlyThrAlaGluArgSerLeuLeu 500
DB 1567 AGCAAGCATATGCAATCGGTGAGAACACAGACAGATGGACCGCCCAAGAGGTCCTTTTAA 1626
QY 501 SerProLeuHisArgSerGlySerValGluAspAsnIleHisThrSerPheLeuPheGly 520
DB 1627 TCTCTCATGATCGAAGTGGAGACGTCGAGAGCAATTAACACACAGGCTCTTTTGGGC 1686
QY 521 LeuSerThrSerGlnGlnHisLeuThrIleYsSerAlaGlyIleGluYsAlaArgIleHis 540
DB 1687 CTTTCCACGACGACAGACCTCAAGAGCTGCTGGCTGGGCTTTAAAGGCTGGCAC 1746
QY 541 SerAspIleLeuAlaProGluThrSerThrProSerLeuThrSerSerThrIleYsPheAla 560
DB 1747 TCGGATATCTTGGCCCCCAGACTCTTACCCCTTCCCTGACCAAGCTGGTATTTTGGC 1806

```

```

QY 561 ThrGluSerSerHisPheIleSerAlaIleIleIleIleIleIleIleIleIleIle 580
DB 1807 ACAGAGTCTCTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1866
QY 581 AlaIleSerCysSerGlnLeuProThrCysGlyAspGluAlaIleYsSerValArgArg 600
DB 1867 GCTTACAGCTGACGACGACCTGCCCACTTGGGAGAACCAAGCTTATCTGTGGCCAGGCG 1926
QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerThrIleGluGluSerProPheGlu 620
DB 1927 CAGAGCCAAAGTACAGAGCTGATCCCGCGAGCTGGGATGAGAGAGAGCCCTTTGAA 1986
QY 621 LysGlnPheLysArgArgSerCysGluMetGluPheGlyIleSerIleMetSerGluAsn 640
DB 1987 AAGCATTTTAAACGCAAGACCTGCAAAATGATTTGAGAGAGCATCATGTCAGAGAAC 2046
QY 641 ArgSerArgGluGluLeuGlyIleYsValGlySerGlnSerSerPheSerGlySerMetGlu 660
DB 2047 AGGTACAGGAGAGAGCTGGGAGAAAGTGGCAGTCAAGTCTTTCGGGACGATGAA 2106
QY 661 ILeIleGluValSer 665
DB 2107 ATCATTCAGGCTTC 2121

RESULT 5
US-09-964-277-1
; Sequence 1, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964,277
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-1

Alignment Scores:
Pred. No.: 0 Length: 3496
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
Gaps: 0

US-10-029-345a-109 (1-665) x US-09-964-277-1 (1-3496)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 562 ATGGCCCAAGATATATGGAATCAATGTTGTTACTBAGAGGTTGGTGCTGCTGGAA 621
QY 21 SerGlyThrGluYsValLeuLeuIleAspSerArgProPheValGluIleYsThrSer 40
DB 622 AGTGAACGGAAAAAGTCTGCTAATTGATAGCGGCACTTGTGGAATAACAATACATCC 681
QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIleYsLeuMetLysArgArgLeuGln 60
DB 682 CACATTTTGAAGGCCATTAATATCAACTGCTCCAAAGCTTATGAGAGGAGAGGTTGCAACG 741
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaIleHisIleYsValAspIleAsp 80
DB 742 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCAAAACATTAAGGTTGACATTGAT 801
QY 81 CysSerGlnLysValValValIleYsAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 802 TGCAGTCAGAAAGGTGTAGTTATGATTAAGATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 861

```

QY 101 AspCybPheLeuThrValLeuLeuGlyValLeuGlyValSerPheLeuSerValHisLeu 120
 DB |||||
 DB 862 GACTGTTTCTCACTGACTTCTGGGTAACCTGAGAAAGACTTCAACTGTTTCACTG 921
 QY 121 LeuAlaGlyPheAlaGluPheSerArgCybPheProGlyLeuCybGluGlyLysSer 140
 DB 922 CTGGCAGGTGGGTTTGCTGAGTCTCTCGTGTTCCTGCGCTCTGTGAGGAAATCC 961
 QY 141 ThrLeuValProThrCysLysSerGlnProCysLeuProValAlaAsnLysGlyProThr 160
 DB 982 ACTCTAGTCCCTCACTGCACTTCTGAGCCCTGCTTACTGTTGCAACATTTGGGCCAAC 1041
 QY 161 ArgLysLeuProAsnLeuThrLeuGlyCysGlnArgAspValLeuAsnLysGlyLeu 180
 DB 1042 GCAATCTCTCCCACTTATCTTGGCTGCGCAGAGATGTCCTCAACAGAGAGCTGATG 1101
 QY 181 GlnGlnAsnGlyLysGlyValLeuAsnLysSerThrCysProLysProAspPhe 200
 DB 1102 CACGAGATGGATTGGTTATGTGTTAAATGCCAGCAATCCTGTCCAAAGCCTGACTTT 1161
 QY 201 LysProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlyLysLeu 220
 DB 1162 ATCCCGAGTCTCATTTCTGCGGTGCTGTGATGACAGCTTTGTGAGAAATTTTG 1221
 QY 221 ProThrLeuAspLysSerValAspPheLysLysValAlaValLysAsnGlyCysVal 240
 DB 1222 CCGTGTGGACAAATAGTAGATTTATGAGAAAGCAAAAGCTCCAAATGATGATGTT 1281
 QY 241 LeuValHisCysLeuAlaGlyLysSerArgSerAlaThrLysAlaLysAlaLysLysLys 260
 DB 1282 CTAGTGACAGTTTAACTGAGGATCTCCGCTCCGCACTGCTGCTGCTGCTGCTGCTG 1341
 QY 261 LysArgMetAspMetSerLeuAspGlyAlaLysArgPheValLysGlyLysValProThr 280
 DB 1342 AAGAGGTGACATGCTTTAGATGAGAGCTTACAGATTTGTGAAGAAAAAAGCACTACT 1401
 QY 281 LysSerProAsnPheAsnPheLeuGlyLysLeuAspThrGlyLysLysLysLysLys 300
 DB 1402 ATATCTCCAACTCAATTTTCTGGGCAACTCTGAGCATGAGAAAGATTTAAGAAC 1461
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLysLysLysLysLysLys 320
 DB 1462 CAGACTGAGCACTCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGAGGAGCAAAAT 1521
 QY 321 GluProValProAlaValSerGlyGlyGlyLysSerGlyLysProLysProPro 340
 DB 1522 GAACGTGCTCCCTGCTGCTCAGAGGGTGAGAGAAAGCAGAGAGCCCTCAGTCCACC 1581
 QY 341 CysAlaAspSerAlaThrSerGlyAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 DB 1582 TGTGCCACTCTGCTACCTCAGAGGCGAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1641
 QY 361 ProSerValProSerValGlnProSerLeuGlyLysAspSerProLeuValGlnAlaLeu 380
 DB 1642 CCGAGCTGCCCGCAGCGGAGCGCTGCTGTAGAGACAGCCCGCTGATCAGCGCCCTC 1701
 QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlyLysAspSerLysLysLysLysLys 400
 DB 1702 AGTGGCTGACCTGCTCCGAGAGAGGCTGGAAGACAGCAATAGCTCAAGCCTTCTTC 1761
 QY 401 SerLeuAspLysLysSerValSerThrSerLysSerMetAlaLysSerLeuHisGlyPhe 420
 DB 1762 TCTCTGATATCAAAATAGTTTCAATATTCAGCGAGCAGTGCAGATCTTCAACATGGCTTC 1821
 QY 421 SerSerSerGluAspAlaLeuGlyLysLysPheSerThrThrLeuAspGlyLysAsn 440
 DB 1822 TCTCTCAACAAAGATGCTTTGGAATATCAAACTTCCACTGCTGATGAGGACCAAC 1881
 QY 441 LysLeuCysGlnPheSerProValGlnGlyLysSerGlyLysLysLysLysLysLys 460
 DB 1882 AAGCTATGACAGTCTCCCTGTTTCAAGAACTATCGAGAGAGACTCCCGAAACAGTCCCT 1941
 QY 461 AspLysGluGluAlaSerLysLysLysLysLysLysLysLysLysLysLysLysLys 480

DB 1942 GATTAAGAGAGAGCCAGCATCCCGAAGAGCTGACACCGCCAGGCTTCAGACAGCAG 2001
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 DB 2002 AGCAAGCGATTGATTCGCTGAGAACACAGACAGAGCGCCCGCAGAGGTCCCTTTTA 2061
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnThrHisPheSerPheLeuPheGly 520
 DB 2062 TCTCCACTCATCGAAGTGGAGCGTGGAGAGCAATTTACACAGCTTCTTTTCGGC 2121
 QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysLysGlyLysPhe 540
 DB 2122 CTTTCCACAGCAGCAGACGACCTCAGAGATGCTGCTGCTGGCCTTAAGGCTGGAC 2181
 QY 541 SerAspLysLeuAlaProGlnThrSerThrProSerLeuThrSerSerThrPheAla 560
 DB 2182 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACGAGCTGTATTTTGGC 2241
 QY 561 ThrGluSerSerHisPheThrSerAlaSerAlaLysThrGlyGlyLysSerLysSer 580
 DB 2242 ACAGAGTCTTCACACTTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2301
 QY 581 AlaLysSerCysSerGlnLeuProThrCysGlyAspGlnValLysSerValArgArgArg 600
 DB 2302 GCCTACAGCTGACGACGACGCTGCCACTTGCAGAGACCAAGTCTATCTGCGCAGGCG 2361
 QY 601 GlnLysProSerAspArgAlaAspSerArgSerThrHisGluGluSerProPheGlu 620
 DB 2362 CAGAAACCAAGTACAGAGCTGACTGCGCGGAGCTGCGCATGAAAGAGCCCTTTGAA 2421
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGlnSerLysMetSerGluAsn 640
 DB 2422 AAGCAGTTTAAACGAGAGAGCTGCAATGAAATTTGGAGAGCATCATGTCAGAGAAC 2481
 QY 641 ArgSerArgGluGlyLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 DB 2482 AGGTACCGGAGAGAGTGGGGAAAGTGGCAGTCACTGATCTTTTCCGCGACAGCATGAA 2541
 QY 661 IleIleGluValSer 665
 DB 2542 ATCATTTAGAGTCTCC 2556

RESULT 6
 US-09-816-494-1
 ; Sequence 1, Application US/09816494
 ; Patent No. US20020034807A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Meyers, Rachel A.
 ; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
 ; FILE REFERENCE: PHOSPHATASE MOLECULES AND USES THEREFOR
 ; CURRENT APPLICATION NUMBER: US/09/816,494
 ; CURRENT FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: US 60/191,858
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3544
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (589) ... (2583)
 US-09-816-494-1

Alignment Scores:
 Pred. No.: 0
 Score: 3406.00
 Percent Similarity: 99.85%
 Best Local Similarity: 99.70%
 Query Match: 99.65%
 Length: 3544
 Matches: 663
 Conservative: 1
 Mismatches: 1
 Indels: 0

341 CysAlAspSerAlaThrGluValIleLeuMet
.....CAGTCCACC 1608

Publication No. US20040009501A1

```

RESULT 7
US-10-377-072-25
; Sequence 25, Application US/10377072
; Publication No. US20040009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc

```

```

APPLICANT: Curtis, Rory A.J.
APPLICANT: Logan, Thomas Joseph
APPLICANT: Gluckemann, Maria A.
APPLICANT: Meyers, Rachel E.
APPLICANT: Williams, Mark J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Miyoung
APPLICANT: Tsai, Fong-Ying
TITLE OF INVENTION: NOVEL 2869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
FILE REFERENCE: MP103-0180MIM
CURRENT APPLICATION NUMBER: US/10/377,072
CURRENT FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: US 09/895,860
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215,370
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/723,806
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 60/187,455
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/843,297
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199,801
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: US 09/861,801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205,508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/816,494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/815,419
PRIOR FILING DATE: 2001-03-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 3544
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (589) ... (2586)
US-10-377-072-25

Alignment Scores:
Pred. No.: 0
Score: 3406.00
Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 99.65%
Length: 3544
Matches: 663
Conservative: 1
Mismatch: 1
Indels: 0
Gaps: 0
US-10-029-345a-109 (1-665) x US-10-377-072-25 (1-3544)
QY 1 MetAlhIeGluNecIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 589 ATGGCCCATGAGATGATTGGAACTTGTACTGAGAGAGTTGGGCTGCTGCGAA 648
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 649 AGTGAAGCGAAAGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 708
QY 41 HisIleLeuGluValAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
Db 709 CACATTGGAAAGCATTAATATCACTGCTCCAGGTTATGAAGGAAAGTTGCAACG 768
QY 61 AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Db 769 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAAACATAGAGTTGAT 828
QY 81 CysSerGlnIleValValValValValValValValValValValValValVal 100

```

```

Db 829 TGCACTCAGAAAGTGTATGATTACATCAAGCTCCAAAGATGTTGCTCTCTCTCA 888
QY 101 AspCyPheLeuThrValLeuLeuGlyIleValLeuGlySerPheAsnSerValHisLeu 120
Db 889 GACTGTTTCTCAGTGTACTTCTGGGTAACTGGAAGAGCTTCAACTCTGTCACTG 948
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgPheProGlyLeuGlyGlySer 140
Db 949 CTTCAGAGTGGTGTCTAGTCTCTCTGTTCTCTGCTGCTCTGCTGCTGCTGCTG 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 1009 ACTTAGTCCCTACCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTC 1068
QY 161 ArgIleLeuProAsnLeuIleGlyCysGlnArgAspValLeuAsnIleGlyLeu 180
Db 1069 CGAATCTTCCCAATCTTATCTTGGCTGCCAGCGAGATGCTCTCAACAGAGCTGATG 1128
QY 181 GlnGlnAsnGlyIleGlyThrValLeuAsnAlaSerThrCysAspProLysProAspPhe 200
Db 1129 CAGCAAGATGGATGGTGTATGTTAATGCGAGCAATACCTGTCCAAAGCTTGACTTT 1188
QY 201 IleProGlySerHisPheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
Db 1189 ATCCCGAGTCTCATTTCTGGGTGTGCTGATGAGACAGCTTTGTGAGAAATTTTG 1248
QY 221 ProThrLeuAspLysSerValAspPheIleGlyValAlaLeuAlaSerAsnGlyCysVal 240
Db 1249 CCGTGGTGTGAACAATCAGTACATTTCAATGGAAGCAAAAGCCCTCAATGATGTGTT 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaThrIleMet 260
Db 1309 CTAGTCACTGTGTAGTGTGAGATCTCCGCTCCGACCATGCTATCCCTACATCAG 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaIleArgPheValIleGlyIleValAspProThr 280
Db 1369 AAGAGATGAGATGCTTATGATGAGCTTCAAGATTTGTGAAGAAAGAAAGAAAGCTACT 1428
QY 281 HisSerProAsnPheAsnPheLeuGlyIleLeuLeuAspThrGlyIleValHisLeuVal 300
Db 1429 ATATCTCCAACTTCAATTTCTGGGCAACTCTGAGTATGAGAAAGATTAAGAAC 1488
QY 301 GlnThrGlyAlaSerGlyProLysSerIleValLeuLeuAspThrGlyIleValHisLeuVal 320
Db 1489 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTCTGACCTGAGAGAGCCAAAT 1548
QY 321 GluProValProAlaValSerGluGlyIleLysSerGluThrProLeuSerProPro 340
Db 1549 GAACCTGTCTCTGCTCTCTCAAGAGGTGACAGAAAGAGAGAGCCCTCACTCAACCC 1608
QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
Db 1609 TGGCGCACTCTGCTACCTCAGAGGCAAGCAAAAGCCCTGTGATCCCGCAAGCTG 1668
QY 361 ProSerValProSerValGlnProSerLeuLeuLeuAspSerProLeuValGlnAlaLeu 380
Db 1669 CCAGAGTCCCAAGGCTGAGCGCTGCTGATGAGAGAGAGCCGCTGTATACAGGCGCTC 1728
QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIleLeuValAspSerPhe 400
Db 1729 AGTGGCTCACTGTCTCCGACAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1788
QY 401 SerLeuAspIleLysSerValSerIleSerAlaSerMetAlaIleSerLeuHisGlyPhe 420
Db 1789 TCTCTGATATCAAAATCAGTTTCATATTCAGCCAGCATGCGCATCTTATACATGCTTC 1848
QY 421 SerSerSerGluAspAlaLeuGlnIleValIleValIleValIleValIleValIleVal 440
Db 1849 TCTCTATCAGAAAGTCTTGAATATCAAACTTCACTACTCTGATGAGAGAGAGAGAG 1908
QY 441 LysLeuCysGlnPheSerProValGlnIleLeuSerGluGlnIleThrProGlyThrSerPro 460

```



```

Db      1909 AACCTATGCCAGTCTCCCTGTTACAGAACTATCGACGACACTCCGGAACCACTGCT 1968
QY      461 AAsPlySGluGluA1SerTLeProLySLeuGlnThra1AarPProSerAAsPserGln 480
Db      1969 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2028
QY      481 SerLySArgLeuH1SerSerValArgThSerSerSerGlyThz1Aa1InArgSerLeu 500
Db      2029 AGCAAGCGATTGCAATTCGGTCAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2088
QY      501 SerProLeuH1SArGSerGlySerValGluAAsPserThiAThSerPheLeuPheGly 520
Db      2089 TCTCCACTGCATCCAAAGTGGAGAGGTGGAGAACAAATACCAACAGCTTCCTTTCCGCG 2148
QY      521 LeuSerThrSerGlnGlnH1SLeuThrLySserA1GlyLeuGlyLeuLySGLYTPPhis 540
Db      2149 CTTTCCACCAAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2208
QY      541 SerAsp1LeuA1AProGlnThrSerThrProSerLeuThrSerSerTrpTyPheA1a 560
Db      2209 TCGGATATCTTGGCCCCCCCCAGACCTCTTACCCCTTCCCTGACCAAGAGCTGGTAATTTGCC 2268
QY      561 ThrGluSerSerHisPheTySerA1SerA1a1eTyrglyGlySerA1aSerTySer 580
Db      2269 ACAGAGTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCT 2328
QY      581 AlaTySerCySerGlnLeuProThrCySerGlyAspGlnValTySerValArgAArg 600
Db      2329 GCCTACAGCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2388
QY      601 GlnLySProSerAAsPArgA1AAsPserAArgTySerTrpHisGluGlnGlnProPheGlu 620
Db      2389 CAGAAAGCCAAAGTGAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2448
QY      621 LysGlnPheLySArgAArgSerCyGlnMerGlnPheGlyGluSerTLeuSerGluAsn 640
Db      2449 AAGCAGTTTAAAGCAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2508
QY      641 ArgSerArgGluGluLeuGlyLySValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db      2509 AGGTACAGGGAGAGAGTGGGGAAAGTGGGACAGTCAAGTCTTTCCGGGAGCATGGA 2568
QY      661 Ile1IleGlnValSer 665
Db      2569 ATCATTTGAGGTCTCC 2583

```

RESULT 8

```

US-10-425-114-26234
; Sequence 26234, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 26234
; LENGTH: 3625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; OTHER INFORMATION: Clone ID: LIB4119-028-H6_FLI
US-10-425-114-26234

```

Alignment Scores: 0 Length: 3625

Score: 3406.00
Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 99.65%
DB: 13
Gaps: 0

US-10-029-345a-109 (1-665) x US-10-425-114-26234 (1-3625)

```

QY      1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
Db      692 ATGGCCCATGAGATGATGGAAGTCACTCAATTTGTAAGAGAGTGGTGGCTGCTGGAA 751
QY      21 SerGlyThrGlyValLeuLeu1AAsPserAArgProPheValGlnTyPheSer 40
Db      752 AGTGAAGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 811
QY      41 HisLeuGlnAla1Aaen1Aaen1Aaen1Aaen1Aaen1Aaen1Aaen1Aaen1Aaen1Aa 871
Db      812 CACATTTTGGAGGCACTTAATATCAATCTCTCAAGCTTATGAAAGAGAGAGAGAGAGAGAG 871
QY      61 AspLySValLeu1eThrGluLeu1eGlnH1SerA1a1eTySLeuValAsp1eAsp 80
Db      872 GACAAAGTGTAAATTCAGAGCTCATCCAGATTCAGGAAACATTAAGTTGACATGAT 931
QY      81 CySerGlnLySValValValTyraSPGlnSerSerGlnAspValA1aSerLeuSer 100
Db      932 TGCAGTCAGAGAGGTGTATTTACATCAAGAGCTCCAAAGTGTGGCTCTCTCTCTTCA 991
QY      101 AspCyPheLeuThValLeuLeuGlyLySLeuGlnLySserPheAsnSerValHisLeu 120
Db      992 GACTGTTTCTCACTGTAATCTTGGGTAACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1051
QY      121 Leu1aGlyGlyPheA1aGluPheSerAArgCyPheProGlyLeuCySerGluGlyLyS 140
Db      1052 CTTGAGAGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1111
QY      141 ThrLeuValProThrCyS1eSerGlnProCySLeuProValA1aAsn1eGlyProThr 160
Db      1112 ACTCAAGTCTTCACTGTAATCTTGGGTAACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1171
QY      161 Arg1IleLeuProAsnLeuTyLeuGlyCyGlnAArgAspValLeuAsnLySLeu1e 180
Db      1172 CGAATCTTCCCAATCTTATCTTGGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1231
QY      181 GlnGlnAsnGlyIleGlyTyValLeuAsnA1aSerTyThrCyPProLyProAspPhe 200
Db      1232 CAGCAGAAATGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1291
QY      201 IleProGluSerHisPheLeuArgValProValAAsnAspPheCySGLYValIleLeu 220
Db      1292 ATCCCGAGTCTCATTTCTGGGTGGCTGTGAGATGACAGCTTTGTGAGAAATTTTG 1351
QY      221 ProTrpLeuAspLySserValAAsPhe1eGlnLySAlaLySAsnAsnGlyCyVal 240
Db      1352 CGGTGGTGGCAAAATCGTGAATTTCAATTGAGAAAGCAAAAGCTTCATGATGATGATGAT 1411
QY      241 LeuValHisCyLeuA1aGly1eSerAArgSerA1aThr1eAla1eAlaTy1eMet 260
Db      1412 CTAGTGCACTGTTAAGTGGGATCTCCCGGCTCCGACCACTCCATACGCTTACATCAG 1471
QY      261 LysArgMetAspMetSerLeuAspGluA1aTyraSPGlnSerPheValIleGlyLySArgProThr 280
Db      1472 AAGAGATGAGACATGCTTTAAGATGAAGTTCAGATTTGTGAAAGAAAGAAAGAAAGAAAG 1531
QY      281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyGlnLySArg1eLySAsn 300
Db      1532 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGGAATTAAGAGAGAGAGAGAGAGAGAGAG 1591
QY      301 GlnThrGlyA1aSerGlyProLySserLySLeuLySLeuGlnLySLeuGlnLySProAsn 320
Db      1592 CAGACTGAGACATCAGAGGCAAGAGCAAACTCAAGCTGCTGCACTGAGAGAGCAAAAT 1651
QY      321 GluProValProAlaValSerGluGlyGlnLySserGluThrProLeuSerProPro 340

```



```

Db      1652 GAACCTCTCCCTGCTGTCTCAGAGGGGTGACAGAAAAGCCGAGACGCCCTCACTGACACCC 1711
Qy      341  CysAlaAapSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
Db      1712 TGTGGCCACTCTGTCTGACTCAGAGGCGAGGAGCAAAAGGCGCGGTGATCCCGCAGCGTG 1771
Qy      361  ProSerValProSerValGlnProSerLeuGlnAapSerProLeuValGlnAlaLeu 380
Db      1772 CCCAGCTGCGCCAGCGAGCGCTGCTGTTAGAGACAGCCCGCTGTACAGGCGCTC 1831
Qy      381  SerGlyLeuHisLeuSerAlaAapArgLeuGlnAapSerAnlyLeuLysValArgSerPhe 400
Db      1832 AGTGGGCTGACCTGTCGCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTCTTC 1891
Qy      401  SerLeuAapLleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
Db      1892 TCTCTGATATCAAAATCAAGTTTCAATATTCAGCCAGCATGCGAGCATCTTACATGCGCTTC 1951
Qy      421  SerSerSerGlnAapAlaLeuGlnTyrTyrLysProSerThrThrLeuAapGlyThrAsn 440
Db      1952 TCTCTCATCAAGAAAGTCTTGAAATCTACAAACCTTCCACTCTGAGATGGGACCAAC 2011
Qy      441  LysLeuCysGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
Db      2012 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATCGAGAGAGACTCCGAAACCAAGTCTC 2071
Qy      461  AapLysGlnGlnAlaSerLleProLysLeuGlnThrAlaArgProSerAapSerGln 480
Db      2072 GATTAAGAGAGAACCCAGCATCCCAAGAAAGCTGAGAGCCCGCAGGCTTACAGACGAC 2131
Qy      481  SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db      2132 AGCAAGGATGTGATTCGTGTAGAACAGACAGCGTGGACCCGCGCAGAGTCTCTTTTA 2191
Qy      501  SerProLeuHisArgSerGlySerValGlnAapAsnTyrHisThrSerPheLeuPheGly 520
Db      2192 TCTCTCACTGATCGAATGGAGAGCGTGGAGAGCAATTACCAACACGACTTCTTTTCGCG 2251
Qy      521  LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheHis 540
Db      2252 CTTTCCACCGACGACGACACTCTCAAGAGTCTGCTGGCGCTTAAGGCGTGGCAC 2311
Qy      541  SerAapLleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrTyrPheAla 560
Db      2312 TCGGATATCTTGCGCCCGCCAGACCTTACCCCTTCTGACACGACGCTGGTATTTTGGCC 2371
Qy      561  ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSer 580
Db      2372 ACAGAGTCTCAACATTTCTACTCTGCTCAGCCATCTACGAGAGGAGTGCAGTACTCT 2431
Qy      581  AlaTyrSerCysSerGlnLeuProThrCysGlyAapGlnValTyrSerValArgArgArg 600
Db      2432 GCCTACAGCTGACGACGACGCTGCCCACTTGCGAGAGCAAGTCTATTTGCGCCAGCGG 2491
Qy      601  GlnLysProSerSerAapArgAlaAapSerArgArgSerTyrPheGlnGlnLysSerProPheGln 620
Db      2492 CAAAGGCAAGTACAGAGGTGACTCCGCGGAGCGTGCATGAAAGAGACCCCTTTGAA 2551
Qy      621  LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerLleMetSerGlnAsn 640
Db      2552 AAGCAGTTTAAAGCAGAGACGCTGCCAATGGAAATTTGGAGAGACATCATGTACAGAAC 2611
Qy      641  ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
Db      2612 AGGTACAGGGAAGAGCTGGGGAAGTGGGAGTCAGTAGCTTTTTCGGGCAACATGAA 2671
Qy      661  IleGlnGlnValSer 665
Db      2672 ATCATTTAGAGTCTTC 2686

```

RESULT 9
US-10-343-357-17

```

; Sequence 17, Application US/10343357
; Publication No. US20040058341A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y.TOM
; APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
; APPLICANT: YAO, Monique G.; BUREFORD, Neil
; APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
; APPLICANT: LEE, Ernestine A.; HARALIA, April J.A.
; APPLICANT: LU, Dying Aina M.; TRIBOULEY, Catherine M.
; APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
; APPLICANT: YUE, Henry; WARREN, Bridget A.
; APPLICANT: NGUYEN, Daniel B.; CHAWLA, Narinder K.
; APPLICANT: KEARNEY, Liam
; TITLE OF INVENTION: PROTEIN PHOSPHATASES
; FILE REFERENCE: FI-0173 PCT
; CURRENT FILING DATE: 2003-01-28
; PRIOR FILING DATE: PCT/US01/23716
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,679
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/223,272
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/224,309
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 60/226,728
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/229,254
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: US 60/231,366
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 3766
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CB1
US-10-343-357-17

Alignment Scores:
Pred. No.: 0 Length: 3766
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: 13 Gaps: 0

US-10-029-345a-109 (1-665) x US-10-343-357-17 (1-3766)
Qy      1  MetAlaHisGluMetLleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
Db      538  ATGGGCCACGACATGATTTGGAATCTCAATTTGTTACTGAGAGGTGGTGGCTGCTGCGAA 597
Qy      21  SerGlyThrGlnLysValLeuLeuIleAapSerArgProPheValGlnTyrAsnThrSer 40
Db      598  AAGTGAACGGAAGAAAGTGTGCTAATTGATTAGCGCGGCACTTTGTGGAATACATATCATCC 657
Qy      41  HisLleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGln 60
Db      658  CACATTTTGAAGCCATTATATCAACTCTCCAACTTATGAAGCAAGTTGCAACAG 717
Qy      61  AapLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAapLleAap 80
Db      718  GACAAAGTGTATTTACAGAGCTCAATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 777
Qy      81  CysSerGlnLysValValValTyrAapGlnSerSerGlnAapValAlaSerLeuSerSer 100
Db      778  TGCAGTCAGAAAGTGTAGTTTACGATCAAAAGTCCCAAGATGTTGCTCTCTCTTCA 837

```

QY	101	AspCySPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeu	120
Db	838	GAcTGTTTCTCACAGTACTCTGGGTAACTGGAGAAAGCTTCAACTGTTCCACCTG	897
QY	121	LeuAlaGlyGlyPheAlaGlyIlePheSerArgCysPheProGlyLeuGlyLysSer	140
Db	898	CTTGCAGGTGGGTTTCTCGAGTTCTCTGTTGTTCCCTGGCGCTCTGGAGAAATCC	957
QY	141	ThrLeuValProThrCysGlyIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	958	ACTCAAGCTCCCTACCTGACATTTCTCAGCTTCTCTTAACCTTTCACATATGGAGCAAC	101
QY	161	ArgIleLeuProAsnLeuLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGlyLeuIle	180
Db	1018	CGAATTCTTCCCAATCTTTATCTTGGCTGGCCAGAGATGCTCCACAAAGAGCTGATG	1077
QY	181	GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe	200
Db	1078	CAGCAAGAAAGGATGGTTATGTGTTAAATCCAGCAATCCGTCCAAAGCTCACTT	1137
QY	201	IleProGlnSerHisPheLeuArgValProValAsnArgSerPheCysGlnLysIleLeu	220
Db	1138	ATCCCGAGCTCATTTCTCGTGGTGGCTGTGAATGACAGCTTTGTGAGAAATTTTG	1197
QY	221	ProThrLeuAspLysSerValAspPheIleGlyLysValAlaLysAlaSerAsnGlyCysVal	240
Db	1198	CCGTGGTGGCAAAATCAGTATTTCAATGAGAAAGCAAAAGCTCCAAAGAGATGCTT	1257
QY	241	IleuValHisCysLeuAlaGlyIleSerArgSerHisThrIleAlaTyrIleMet	260
Db	1258	CTAAGTGACCTGTTTAGCTGGAGATCTCCGCTCCGACCATCGCATGCTCATCATCATG	1317
QY	261	LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlyLysArgProThr	280
Db	1318	AAGAGATGAGCATGCTCTTATGATGAAGCTTACAGTTTGTCAAGAAAGAAAGAAAGCTCACT	1377
QY	281	IleSerProAsnPheAsnPheLeuGlyGlyLeuLeuAspTyrGlyLysIleLysAsn	300
Db	1378	ATATCTCCAATCTCAATTTTCTGGCCAACTCTCGACATATGAGAAAGATTTAAGAAC	1437
QY	301	GlnThrGlyAlaSerGlyProLysSerLysLysLeuLeuHisLeuGlyLysProAsn	320
Db	1438	CAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTCTGCACTGGAGAGGCCAAT	1497
QY	321	GlnProValProAlaValSerGlnGlyGlyLysSerGlnThrProLeuSerProPro	340
Db	1498	GAACTGTCCTCGCTGCTCTCAGAGGGTGGACAGAAAAAGCAGACGCCCCAGCTCAACC	1557
QY	341	CysAlaAspSerHisThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal	360
Db	1558	TGTGCCGACTCTGTAACCTCAGAGGACGACGAGCAAAAGCCCTGTGATCCGCAAGCGTG	1617
QY	361	ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu	380
Db	1618	CCCAACGTGGCCAGAGGTGACGCGTGCCTGTTAGAAGACAGCCCGCTGTACAGGCGCTC	1677
QY	381	SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuValArgSerPhe	400
Db	1678	AGTGGACTGCACTGTCCGACAGACAGGCTGGAAGACACAGATTAAGCTCAAGCTTCTTC	1737
QY	401	SerLeuAspLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe	420
Db	1738	TCTCTGGATATCAAAATCAATTCATATTCAGCCACATGAGCAGCATCTTACATAGCTTC	1797
QY	421	SerSerSerGlnAspAlaLeuGlnTyrTyrLysProSerThrThrLeuAspGlyThrAsn	440
Db	1798	TCTCTCAACGAAATATCTTTGGAAATCAACAACCTTCCACTACTCTGGATGGAGCAAC	1857
QY	441	LysLeuGlyGlnPheSerProValGlnLysLeuSerGlnGlnThrProGlnLysSerPro	460
Db	1858	AAACGTATGCAGATTCTCCCTCTTTCAAGAACTATTCGAGAGACGCTCCGAAACAGCTCT	1917
QY	461	AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerArgSerGln	480

Query	Database	Accession	Score	E-value	Length	Mismatches	Conservative	Best Local Similarity	Query Match	DB
QY	Db	1918 GATTAAGAGAGAACCATGATCCCAAGAACTGACAGCCGACGGCTTCAGACAGCCAG	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	481 Ser1ySArGLeuH1sSerValArgThrSerSerSerG1yThr1aG1nA1nSerLeuLeu	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	1978 AGCAAGCATTTGATTCGGTCAGAACCAAGACAGCATGGACCCGCCAGAGATCCCTTTTA	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	501 SerProLeuH1sA1gSerG1ySerValG1uAAspAntYrH1sThrSerPheLeuPheG1y	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	2038 TCTCCATCGATCGAATGAGAGGTGGAGAGAAATTTACCACTCCAGCTTCCTTTTGGC	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	521 LeuSerThrSerG1nG1nH1sLeuThr1ySArG1yLeuG1yLeuY1yG1yYrPh1s	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	2098 CTTTCACACAGCCAGACAGACCTCCAGAAAGTCTGCTGGCTTGGCTTTAAGGACTGGCAC	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	541 SerAsp11eLeuA1aProG1nThrSerThrProSerLeuThrSerSerTrpYrPheA1a	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	2158 TCGAATATCTTGGCCCCCAGACCTTCACCTTCCTCCAGACAGACGCGATATTTTGGC	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	561 ThrG1uSerSerH1sPheYrSer1aSerA1a1eYrG1yG1ySer1aSer1ySer	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	2218 ACGAGATCTTCACCTTCTACTCTCCCTCCAGCATCTACGAGGACAGTGCAGTACTCT	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	581 A1aYrSerYrSerG1nLeuProThrYrG1yS1aSpG1nV1yYrSerV1aYrG1yArg	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	2278 GCTTACACTGACGACGACCTCCCACTGGCGGAAACCAAGCTATCTGTGCGAGGCGG	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	601 G1nYrProSerAspArgA1aAspSerArgYrSerTrpH1sG1uG1uSerProPheG1u	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	2338 CAGAAAGCCAACTGACAGACGCTGACTCCGCGGAGAGCTGGCATAGAGAGACCCCTTGA	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	621 LysG1nPhelysA1aG1ySerYrG1nMetG1uPheG1yG1uSer11eMetSerG1uA	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	2398 AAGCATTTTAAACGACAAACCTCCAAATGAATTTGAGAGAGCATATTCAGAGAAC	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	641 ArgSerArgG1uG1uLeuG1yLysValG1ySerG1nSerPheSerG1ySerMetG1u	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	2458 AGGTCAAGGAGAAAGCTGGGAAAGTGGCAGTCACTTACGCTTTTGGGAGCATGANA	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	661 11e11eG1nV1aSer 665	3406.00	4790	663	1	1	99.704	99.658	17
QY	Db	2518 ATCATTGAGCTCTCC 2532	3406.00	4790	663	1	1	99.704	99.658	17

US-10-029-345a-109 (1-665) x US-10-648-593-115 (1-4790)

QY 1 MetAlHISGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 Db ATGGCCCATGAGATGATGGAACTCAATGTGTAATGAGAGAGTGGTGGCTCGCTGGAA 243
 QY 21 SerGIYThrGluIleValLeuLeuIleAspSerArgProPheValGluIYrAsnThrSer 40
 Db AGTGGAAAGGAAAGAGTGGCTTAATGATGAGCGGCAATTTGGAAATCAATACATCC 303
 QY 41 HisIleLeuGlnIleAlaIleAsnIleAsnCySerIleLeuMetIleAspArgLeuGln 60
 Db CACATTTTGGAGCCATTAAATCACTGCTCAAGCTTATGAAAGGAAAGTTGCAACAG 363
 QY 61 AspIleValIleuIleThrGluLeuIleGlnHisSerAlaIleHisIleValIleAsp 80
 Db GACAAAGTGTAAATTCAGAGCTCATCCAGCATTCAGCCAAACATAAGGTGACATGGAT 423
 QY 81 CySerGlnIleValIleValIleYrAspGlnSerSerGlnAspValAlaSerIleSer 100
 Db TGCAGTCAGAAAGTGTAGTTAGATCAAAAGCTCCCAAGATGTCCCTCTCTCTCA 483
 QY 101 AspCyPheLeuThrValIleuLeuGlyIleuGlyIleuSerPheAsnSerValHisLeu 120
 Db GACTGTTTCTCACTGTAATTCGGGTAAATCGAAGAAAGCTTCACTCTGTCACCTG 543
 QY 121 LeuAlaGlyIleuPheAlaGluPheSerArgCyPheProGlyLeuCyGluGlyIleuSer 140
 Db CTTCGAGGTGGTGTCTGAGTTCCTCGTGTGTTCTCGCTCTGTCAGAAAGTATCC 603
 QY 141 ThrIleValProThrCyValIleSerGlnProCyLeuProValAlaAsnIleGlyProThr 160
 Db ACTCATGCTCCATCCGCACTTCACCTTCGCTTACCTGTGCAACATGGGCAACCC 663
 QY 161 ArgIleLeuProAsnLeuIYrLeuGlyCyGlnArgAspValIleuAsnIleuGluLeu 180
 Db CGAATTCCTCCCAATCTTATCTTGCTGCAAGAGATGTCTCAACAAAGAGCTGATG 723
 QY 181 GlnGlnAsnGlyIleGlyIYrValIleuAsnAlaSerIYrThrCyProIleYrAspPhe 200
 Db CAGCAAAATGGGATGTGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTT 783
 QY 201 IleProGlnSerIlePheLeuArgValProValAsnAspSerPheCyGluIleu 220
 Db ATCCCGAGCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAAATTTTG 843
 QY 221 ProTrpLeuAspIleSerValAspPheIleGlyIleValAlaIleAsnGlyCyVal 240
 Db CCTGTGTGGCAAAATCAGTATGATTCATTGAGAAAGCAAAAGCTTCAATGATGTAT 903
 QY 241 LeuValHisCyLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
 Db CTAGTCACTGTGTTACTGGGATCTCCGCTCGCACCATCTGCTATGCTACATCATG 963
 QY 261 IYrAspMetAspMetSerIleuAspGlnIleYrArgPheValIleGluIleuAspProThr 280
 Db AAGAGGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAACCTTACT 1023
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIYrGluIleuIleuAsn 300
 Db ATATCTCCAAATCTCAATTTCTGGGCCAATCTCTGACATATGAAAGAAATTAAGAAC 1083
 QY 301 GlnThrGlyAlaSerGlyIleuIleuSerIleuIleuIleuIleuIleuIleuIleu 320
 Db CAGACTGAGACATCAAGGCGCAAGCAAACTCAACTGCTGCACTGGAGAAAGCAAAAT 1143
 QY 321 GluProValProIleValIleSerGluIleuIleuIleuIleuIleuIleuIleuIleu 340
 Db GAACCTGTCTCTGCTGTCAGAGGGTGGCAAGAAAGCAAGAGCGCCCTCACTCAACCC 1203
 QY 341 CyAlaAspSerAlaThrSerGlnAlaAlaGlyIleuIleuIleuIleuIleuIleuIleu 360
 Db TGTGCCGACTCTGCTACTCAGAGGCAAGCAAGCAAAAGCCCTGTGATCCCGCAAGCTG 1263

QY 361 ProSerValProSerValGlnProSerLeuGluAspSerProIleValGlnAlaLeu 380
 Db CCAAGCTGCCCAAGCTGAGCGCTGCTGTGTAAGAGAACCGCGTGGTACAGGGCTC 1323
 QY 381 SerGIYLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIleuIleuIleuIleu 400
 Db AGTGGCTCACCTGTGCCCAAGAGCTGGAAGAACAGCAATAGCTTCAAGCTTCTTC 1383
 QY 401 SerLeuAspIleuIleuSerValSerIYrSerAlaSerMetAlaAlaSerIleuHisGlyPhe 420
 Db TCTCTGATATCAAAATCAGATTCAATTCAGAGCAGCATGGAGCATCTTACATGCTTC 1443
 QY 421 SerSerSerGlnAspAlaLeuGlnIleuIYrIYrIYrProSerThrThrIleuAspGlyThrAsn 440
 Db TCTCATCGAAGATGCTTGAATATCAAACTTCACTTCACTTGAATGGAGCAAC 1503
 QY 441 IYrLeuCyGlnPheSerProValGlnGluLeuSerGluGlnIleuProGluIleuThrSerPro 460
 Db AAGCTATGCCAGTTCCTCCCTGTCAGAACTATCGGAGCAAGCTCCGAAACCAATCCT 1563
 QY 461 AspIleGluGlnAlaSerIleProIleuIleuIleuIleuIleuIleuIleuIleu 480
 Db GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGACAGCCGAGCTTCAGACAGCCAG 1623
 QY 481 SerIYrArgLeuHisSerValArgThrSerSerSerGlyIleuIleuIleuIleuIleu 500
 Db ACGAAGCGATTCATTCGCTGCAAGAACACAGACAGAGGAGCCGCCAGAGGCTCTTTTA 1683
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnIYrHisThrSerPheLeuPheGly 520
 Db TCTCCACTCATCGAAGTGGAGCGTGGAGAGCAATACACACAGCTTCCTTTTCGGC 1743
 QY 521 LeuSerThrSerGlnGlnHisLeuThrIYrSerAlaGlyIleuGlyIleuIleuIleu 540
 Db CTTTCACACAGCAGCAGACCTCACAGAGTCTGCTGGCTGGCTTAAGGCTGGCAC 1803
 QY 541 SerAspIleuAlaProGlnIleuThrSerThrProSerLeuThrSerSerTrpIYrPheAla 560
 Db TCGGATATCTGGCCCCCAGACCTTACCCCTCTCCCTTACCAAGAGCTGTATTTTGCC 1863
 QY 561 ThrGluSerSerHisPheIYrSerAlaSerAlaIleIYrGlyIleuIleuIleuIleu 580
 Db ACGAGTCTCTCACTTCACTCTGCTCAAGCATTACAGAGCGAGTCCAGTTACTCT 1923
 QY 581 AlaIYrSerCySerGlnLeuProThrCyGlyAspGlnValIYrSerValArgArgArg 600
 Db GCTACAGCTGACAGCAGCTGCCCACTTTCGGAGACCAAGTCTATTCTGTGCGCAGCGG 1983
 QY 601 GlnIYrProSerAspArgAlaAspSerArgArgSerTrpHisGluIleuSerProPheGlu 620
 Db CAGAAAGCCAAAGTGAAGAGCTGACCTGCGCGAGAGCTGCAATGAAAGAGAGCCCTTTGA 2043
 QY 621 IYrGlnPheIYrArgArgSerCyGlnMetGluPheGlyGluSerIleuMetSerGluAsn 640
 Db AAGCACTTTAAACGAGAGCTGCCAAATGGAATTTGGAGAGCATCTATGTAGAGAAC 2103
 QY 641 ArgSerArgGluGluIleuGlyIYrValGlySerGlnIleuSerPheSerGlySerMetGlu 660
 Db AGGTCAAGGAAAGAGCTGGGAAAGTGGGCAAGTCAAGTCTTTCTGGGAGCATGAA 2163
 QY 661 IleIleGluValSer 665
 Db ATCATTTGAGGTCTCC 2178

RESULT 11
 US-10-094-749-673
 ; Sequence 673, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI

```
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, KYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOTYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
FILE OF INVENTION: NOVEL FULL-LENGTH cDNA
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 673
LENGTH: 2102
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-673
```

Alignment Scores:

```
Pred. No.: 0
Score: 3399.00
Percent Similarity: 99.70%
Best Local Similarity: 99.55%
Query Match: 16
Matches: 662
Conservative: 1
Mismatch: 2
Indels: 0
Gaps: 0
```

US-10-029-345A-109 (1-665) x US-10-094-749-673 (1-2102)

```
QY 1 MetLAlHISGLMELIIEGLYThrgInIleValThrguAglLeuValAlaleuIeuglu
Db 56 ATGCCCATGATGATGATGGAATCTCAATTTGCTGAGAGGTGGTCTCTCGAA
QY 21 SerGIYThrgInuYValIleuIleuIleAspSerArgProPheValGluTYAsnThrSer
Db 116 AGTGAACCGAAAGAGTCTGCTAATTGATGCGGCATTGGGAATACATACATCC
QY 41 HisIleIeugluAlAlIeAnIleAnIleAnCysSerIysIeuWetIysArgIeugIn
Db 176 CACATTTTGGAGCCATTATATCACTCCAGCTTATGAGCGAAGGTTGCAACAG
QY 61 AspIysValIleuIleThrgIleuIleGlnHisSerAlaIysHisIysValAspIleAsp
Db 236 GACAAAGCTTATATACAGAGCTCATCCAGCATTCACGAATACATAGTTGACATTGAT
QY 81 CysSerGlnIysValIValIValTYAspGlnSerSerGlnAspValAlaserIeuSerSer
Db 295 26 TGCAGTCAGAGGTTGATGATTACATCAAACTCCCAAGATGTGCTCTCTCTTCA
QY 101 AspCysPheIeuThrValIleuIeugIlyIysIeugIlyIysSerPheAsnSerValHisIeu
Db 120 356 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGAGCTTCACTCTTCACTG
QY 121 LeuAlaGlyIyPheAlaGlnPheSerArgCysPheProGlyLeuCysGluGlyIysSer
Db 140 416 CTTCAGAGTGGTGGTGGTGGTCTCTCTGTTTCTCTGAGCTCTTGTGAAGGAAATCC
QY 141 ThrIeuValProThrCysIleSerGlnProCysIeuProValAlaAsnIleGlyProThr
Db 160 476 ACTGATCCCTACCTGACATTTCTCAGCCTTGCTTACCTGTTGCAACATTGGGCCAAC
QY 535
```

```
QY 161 ArgIleIeuProAsnIeuTYIleuGlyCysGlnArgAspValIleuAnIysGluIleu
Db 536 CTAATCTTCCCAATCTTTATTTTCTGCTGAGGAGAGATGCTTCAACAGAGCTGATG
QY 181 GlnGlnAsnGlyIleGlyTYrValIleuAnIaserTYrThrCysProIysProAspPhe
Db 595 596 CACAGAAATGGATGGATTGATTGTTAAATCCAGCAATACCTGTCCAAAGCTTCACTT
QY 201 IleProGlnSerHisPheIeuArgValProValAsnAspSerPheCysGluIysIleu
Db 220 656 ATCCCGAGTCTCAATTTCTCGTCTGCTGCTGAAATGACGCTTTGTGTAAGAAATTTG
QY 221 ProTrpIeuAspIysSerValAspPheIleGlyIysAlaIysAlaserAnGlyCysVal
Db 240 716 CCGTGTGGACCAATAGATGATTCATTGAGAAAGCAAAAGCTCCCAATGATGATG
QY 241 LeuValHisCysIeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTYrIleMet
Db 260 776 CTAGTGCACCTGTTAGTGGATCTCCCGCTCGGACCATGCTATGCGCTACATCATG
QY 261 LysArgMetAspMetSerIleuAspGlnAlaTYrArgPheValIysGluIysArgProThr
Db 280 836 AAGAGATGGACATGCTTTAGATGAGCTTACAGATTGTGAAAGAAAGAAAGCTTACT
QY 281 IleSerProAsnPheAsnIleuGlyGlnIleuIeuAspTYrGluIysIleIysAsn
Db 300 896 ATATCTCCAAACTTCAATTTCTGGCCAACTCTGACCTATGAGAAAGATTAAGAAC
QY 301 GlnThrGlyAlaSerGlyProIysSerIysIeuIysIeuIysGluIysProAsn
Db 320 956 CAGACTGGAGCATTCAGGCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG
QY 321 GluProValProAlaValSerGluGlyGlyIysSerGlnThrProIeuSerProPro
Db 340 1016 GAACCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
QY 341 CysAlaAspSerAlaThrSerGlnAlaIleGlyGlnArgProValHisProAlaSerVal
Db 360 1076 TGTGCGCACTCTGCTTACCTCAGAGCGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG
QY 361 ProSerValProSerValGlnProSerIleuIeuIysAspSerProIeuValGlnAlaIeu
Db 380 1136 CCCAGCGTCCGCGAGCGGAGCGGCTGCTTGAAGAGAACCCCTGCTACAGCGCTTC
QY 381 SerGlyIeuHisIeuSerAlaAspArgIeuIysAspSerAsnIysIeuIysArgSerPhe
Db 400 1196 AGTGGCTGCACTGCTCCGACAGGCTGGAAGACACGATTAAGCTCAAGCGTTCTTC
QY 401 SerIeuAspIleIysSerValSerTYrSerAlaSerMetAlaIaserIeuHisGlyPhe
Db 420 1256 TCTCTGATATCAAAATCACTTTCATTTACGCGAGCATGGCAAGATCTTACATAGGCTTC
QY 421 SerSerSerGlnAspAlaIeugIlyTYrIysProSerThrThrIeuAspGlyTYrAsn
Db 440 1316 TCTCATACGAAGATGCTTTGGAAATATCAAAACCTTCACTGATGGAGCAAC
QY 441 IysIeuCysGlnPheSerProValGlnIleuSerGlnGlnThrProGlnThrSerPro
Db 460 1376 AAGCTATGCAAGTCTCCCTGTTCAAGAACTATCGAGAGAGACTCCCGAAACCAATGCT
QY 461 AspIysGlnGluAlaserIleProIysIleuGlnThrAlaArgProSerAspSerGln
Db 480 1436 GATTAAGAGGAAGCAGAGCTCCCAAGAACTGACAGCTCCAGGCTTCAAGAGCCAG
QY 481 SerIysArgIeuHisSerValArgThrSerSerSerGlyTYrAlaGlnArgSerIleu
Db 500 1496 AGCAAGGCAATGCTTGTGTCAGAACAGAGAGAGGAGAGAGAGAGAGAGAGAGAGAG
QY 501 SerProIeuHisArgSerGlySerValGlnAspAsnTYrHisThrSerPheIeuPheGly
Db 520 1556 TCTCCATGATGAACTGGAGCGTGGAGGACAAATTCACACAGAGCTTCTTTTGGGC
QY 521 IeuSerThrSerGlnGlnIleuThrIysSerAlaGlyIleuGlyIleuIysGlyTYrPheHis
Db 540
```

Db 1616 CTTTCCACCCAGCAGACCTCAGCACTGCTGCGCTTAAAGGCTGGCAC 1675
Qy 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpYrPheAla 560
Db 1676 TCGGATATCTTGAGCCCCCAGACCTCTACCCCTTCCCTGACAGAGCTGGTATTTGGCC 1735
Qy 561 ThrGlnSerSerIshPheYrSerAlaSerAlaIleYrGlyGlySerAlaSerYrSer 580
Db 1736 ACGAGATCTCCACACTTCTACTGCTGCTCAGCCATCAGAGAGAGAGTGCAGTACTCT 1795
Qy 581 AlaYrSerCySerGlnLeuProThrCyGlyAspGlnValYrSerValAlaGlyArg 600
Db 1796 GCCTACAGCTGCGACCGACGCTGCCACTTGGGAGACCAAGTCTATTCTGCGCAGGCG 1855
Qy 601 GlnLysProSerAspAlaGlyAspSerArgYrSerTrpHisGlnLysSerProPheGlu 620
Db 1856 CAGAGCCAAAGTACAGAGCTGACTCGCGCGAGCTGGCATGAGAGAGAGCCCTTGA 1915
Qy 621 LysGlnPheLysArgYrSerCyGlnMetGlnPheGlyGlnSerIleMetSerGluAsn 640
Db 1916 AACGAGTTTAAACGAGAGCTGCCAAATGGAAATTGGAGAGAGCATCATGTCAGAGAAC 1975
Qy 641 ArgSerArgGlnGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db 1976 AGGTACGGGAAGAGCTGGGGAAGTGGGAGTCAAGTCTAGCTTTTCCGAGCAGATGAA 2035
Qy 661 IleIleGlnValSer 665
Db 2036 ATCATGAGGTCTCC 2050
RESULT 12
US-10-296-115-520 ; Sequence 520, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseng Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 520
; LENGTH: 2966
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2966)
; OTHER INFORMATION: n = a, t, c or g
US-10-296-115-520
Alignment Scores:
Pred. No.: 0 Length: 2966
Score: 3398.00 Matches: 662
Percent Similarity: 99.70% Conservative: 1
Best Local Similarity: 99.55% Mismatches: 2
Query Match: 99.41% Indels: 0
DB: 13 Gaps: 0
US-10-029-345A-109 (1-665) x US-10-296-115-520 (1-2966)
Qy 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
Db 23 ATGGCCCATAGATGATGGAACTCAATTTGTTACTGAGAGAGGAGGCTGCTGGAGAA 82
Qy 21 SerGlyThrGlnLysValIleLeuLeuIleAspSerArgProPheValGlnYrAsnThrSer 40
Db 83 AGTGAACGGAAGAAAGTGTCTGTAATTGATAGCCGGGCAATTGTGGAAATACATATCC 142

Qy 41 HisIleLeuGlnAlaIleAsnIleAsnCySerIleLeuMetLysArgLeuGln 60
Db 143 CACATTTGGAAGCCATTATATCACTGCTCCACCTTATGAGAGGAGTTGCAACAG 202
Qy 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 203 GACAAAGTTTAAATACAGAGCTCATCCAGCAATTCAGGAAACATAGGTGACATTGAT 262
Qy 81 CySerGlnLysValValIleYrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 263 TGCAGTCAGAGAGTTGATTTTACATCAAGCTCCCAAGATGTTCTCTCTCTTCA 322
Qy 101 AspCyPheLeuThrAlaLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db 323 GACTGTTTCTCATGTCATCTTCTGGGTAACTGGAAAGAGCTTCAACTGTTCACCTG 382
Qy 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCyPheProGlyLeuCyGlnGlyLysSer 140
Db 383 CTTCGAGGTGGGTTCCTAGTCTCTCTGTTTCCCTGGCCTCTGTGAAGGAAATCC 442
Qy 141 ThrLeuValProThrCySillSerGlnProCyLeuProValAlaAsnIleGlyProThr 160
Db 443 ACTCTAGTCCCTACCTGCACTTCTCAGCCTTGCTTACTGTTGCCAAACATTTGGGCAACC 502
Qy 161 ArgIleLeuProAsnLeuYrLeuGlyCyGlnArgAspValLeuAsnLysGlnLeuIle 180
Db 503 CGAATTTCTCCCAATTTTATCTTGGCTGCCAGCGAAGTCTCTCAACAGAGGCTGATG 562
Qy 181 GlnGlnAsnGlyIleGlyYrValIleAsnAlaSerYrThrCySProLysProAspPhe 200
Db 563 CAGCAGAAATGGATGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 622
Qy 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCyGlnLysIleLeu 220
Db 623 ATCCCGAGTCTCATTTCTGCGGTGCGCTGCAAGAGACAGCTTTGTGTGAGAAATTTTG 682
Qy 221 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAsnGlyCySVal 240
Db 683 CCGTGGTTGGACAAACAGTAGATTTTCATTGAGAAAGCAAAAGCTTCCAAATGGATGTGT 742
Qy 241 LeuValHisCyLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaYrIleMet 260
Db 743 CTAGTCACTGTTTACCTGGATCTCCGCTCCGCAACATGCTATCCGCTACATCATG 802
Qy 261 LysArgMetAspMetSerLeuAspGlnAlaYrArgPheValLysGlnLysArgProThr 280
Db 803 AAGAGATGAGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAAGACCTACT 862
Qy 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysIleYrAsn 300
Db 863 ATATCTCCAAACTCTCAATTTTCTGGGCAACCTCTGAGCTATGAGAAAGATTAAGAAC 922
Qy 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlnLysProAsn 320
Db 923 CAGACTGGAGCATCAGGGGCCAAAGAGCAACTCAACTGCTGACCTGGAGAGCCAAAT 982
Qy 321 GluProValProAlaValSerGlyGlyGlnLysSerGlnThrProLeuSerProPro 340
Db 983 GAACTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAGAGAGAGCGCCCTCAGTCCACCC 1042
Qy 341 CyValAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerVal 360
Db 1043 TGTGCCGACTGTGCTACCTCAGAGGAGAGAGCAAAAGGCCGTGTGATCCCGCACGCTG 1102
Qy 361 ProSerValProSerValGlnProSerLeuLeuGlnLysSerProLeuValGlnAlaLeu 380
Db 1103 CCCAGCTGCTCCAGCTGAGCGCTGCTGTTAAGAGAGAGCCCGCTGTGATACAGGCGCTC 1162
Qy 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 400
Db 1163 AGTGGGCTGCACTGTCCGAGCAGAGCTGGAAGAGAGCAATAGACTCAAGCTTCTTCTTC 1222

```

QY 401 SerLeuAsp11eLysSerValSerTyrSerAlaSerMetAla1aSerLeuH1sgLypHe 420
DB 1223 TCTCTGGATATCAATCAATCAATTCATATTCACCGACATGGCAGCATCTCTTAATGGCTTC 1282
QY 421 SerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAan 440
DB 1283 TCCCATCAAGAAAGAGCTTGGAAATCAACAACCTTCCACTCTCTGAGTGGACCAAC 1342
QY 441 LysLeuCySGlnPheSerProValGlnGluLeuSerGlnGlnThrAlaArgProSerAspSerGln 480
DB 1343 AAGCTATGCCAGTTCTCCCTGTTTCAGAAACTATCGAGACAGACTCCGAAACAGTCTCT 1402
QY 461 AspLysGluGluAlaSerTyrProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
DB 1403 GATTAAGAGAGAGACACATCCCAAGAGCTCCACACCGGCTTTCAGACAGCCAG 1462
QY 481 SerLysArgLeuH1sSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
DB 1463 AGCAAGCGATTGCTTGGTCAAGAACACAGACAGAGGCAACCGCCCAAGAGTCTCTTTA 1522
QY 501 SerProLeuH1sArgSerGlySerValGluAspAsnTyrH1sThrSerPheLeuPheGly 520
DB 1523 TCTTCACTGATCGAAGTGGAGCGTGGAGACATTCACACACAGCTTCTTTTGGC 1582
QY 521 LeuSerThrSerGlnGlnH1sLeuThrLysSerAlaGlyLeuGlyLysGlyTyrPheH1s 540
DB 1583 CTTTCCACAGCAGCAGACGACTTCAAGAACTCTGCTGGCTGGGCTTAAAGGCTGGAC 1642
QY 541 SerAsp11eLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
DB 1643 TCGGATATCTTGGCCCCCAACCTTACCCCTTCCCTGACACAGCTGGATTTTGGC 1702
QY 561 ThrGluSerSerH1sPheTyrSerAlaSerAlaLeuTyrGlyLysSerAlaSerTyrSer 580
DB 1703 ACAGAGTCTCTCACTTACTCTGCTCTGACCACTACGAGAGCGAGTCCAGTACTCT 1762
QY 581 AlaTyrSerCySGlnGlnLeuProThrCySGlyAspGlnValTyrSerValArgArgArg 600
DB 1763 GCCTACAGCTGACACCGCTGCCCACTTGGCGAGACCAAGTATCTTCTGGCGAGGCGG 1822
QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrH1sGlnGluSerProPheGlu 620
DB 1823 CAGAAAGCAAGTACAGAGCTGACTCGCGCGAGTGGACGAAAGAGAGCCCTTTGAA 1882
QY 621 LysGlnPheLysArgArgSerCySGlnMetGlnPheGlyGluSerTyrLeuMetSerGluAsn 640
DB 1883 AAGCAAGTTTAAAGCGAAGAGCTGCCAATGAAATTTGAGAGAGCATATCTCAGAGAAC 1942
QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerPheSerGlySerMetGlu 660
DB 1943 AGGTCAACGGGAAGAGCTGGGGAAGTGGGACGTCACTTCTTTGGGCGACATGAA 2002
QY 661 Ile11eGluValSer 665
DB 2003 ATCATTTAGAGTCTCC 2017

```

RESULT 13

```

US-10-072-012-257
Sequence 257, Application US/10072012
Publication No. US20040033493A1

```

GENERAL INFORMATION:

```

APPLICANT: Tcherny, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zetunuen, Bryan
APPLICANT: Patrunajan, Meera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Reha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie

```

```

APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimír Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Rieger, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
PRIOR APPLICATION NUMBER: 2002-01-31
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
Remainder Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 257
LENGTH: 2071
TYPE: DNA
ORGANISM: Homo sapiens
US-10-072-012-257
Alignment Scores:
Pred. No.: 0
Score: 3379.50
Percent Similarity: 99.40%
Best Local Similarity: 99.25%
Query Match: 98.87%
DB: 13
Gaps: 1
US-10-029-345a-109 (1-665) x US-10-072-012-257 (1-2071)
QY 1 MetAlaH1sGluMetT1eGlyThrGln11eValThrGluArgLeuValAlaLeuGlu 20
DB 61 ATGGCCCATGAGATATTTGAATCTCAATTTGTACTAGAGCTTGGTGGCTGCTGGAA 120
QY 21 SerGlyThrGluLysValLeuLeu1eAspSerArgProPheValGluTyrAsnThrSer 40
DB 121 AGTGAAGCGAAAGAGCTCTCTAATTTGATGAGCGGCAATTTGTGAAATACATACATCC 180
QY 41 HisLeuGluGluAla1eAsn1eAsnCySerLysLeuMetLysArgArgLeuGlnGln 60
DB 181 CACATTTTGAAGCATTAATATCACTGCTCAACCTTATGAGCGAAGGTTGCAACAG 240
QY 61 AspLysValLeu1eThrGluLeu1eGlnH1sSerAlaLysH1sLysValAsp11eAsp 80
DB 241 GACAAAGTTTATTTACAGAGCTCATTCAGCATTCAGCAACCAATAGTTGACATTTGAT 300
QY 81 CySerGlnLysValAlaValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 301 TGCAGTCGAAGAGTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 360

```


QY 101 AspCyPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheLeuSerValHisLeu 120
DB 361 GACTGTTTCTCACTGACTTCTGGGTAACTGGAGAAAGAGCTTCACTCTGTTCACTTG 420
QY 121 LeuAlaGlyLysPheAlaGlyPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
DB 421 CTTCAGAGTGGGTTTCTGAGTTCTCTGTTGTTTCCCTGCTCTGTAGAGAAATTC 480
QY 141 ThrLeuValProThrCysHisSerGlnProCysLeuProValAlaAsnHisGlyProThr 160
DB 481 ACTAGTCCCTCACTGACTTCTCAGACCTTGCTTACTGCTGCAACATTTGGCCCAAC 540
QY 161 ArgHisLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuLeuGlyLeuHis 180
DB 541 CGAATCTTCCCAATCTTATCTTGCTGCACGAGAAATGCTTCAACAGAGAGCTGATG 600
QY 181 GlnGlnAsnGlyHisGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
DB 601 CACGAGAAATGGGATTTGGTTATGTTTAAATGCCAGCAATCTGCTCAAGAGCTGACTTT 660
QY 201 HisProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlyLysLeu 220
DB 661 ATCCCGAGTCAATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 720
QY 221 ProTyrLeuAspLysSerValAspPheHisGlyLysAlaLysAlaSerAsnGlyCysVal 240
DB 721 CCTGTGTGACAAATCACTAGATTTCTTATGAGAAAGCAAAAGCTTCAATGATGTGTT 780
QY 241 LeuValHisCysLeuAlaGlyLysSerArgSerAlaThrHisLeuAlaHisTyrHisMet 260
DB 781 CTATGTCACCTTTTATGCTGGAATCTCCGCTCCGCCACATCTGCTATGCTTACATCATG 840
QY 261 LysArgMetAspMetSerLeuAspGlyAlaTyrArgPheValHisGlyLysArgProThr 280
DB 841 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGAGCTACT 900
QY 281 HisSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLysHisHisLeu 300
DB 901 ATATCTTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTAAGAAC 960
QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisGlyGlyLysProAsn 320
DB 961 CAGACTGGAGCATCAGGGCCAAAGAGCAATCAAGCTGCTGACCTGGAGAAAGCAAT 1020
QY 321 GluProValProHisValSerGlyGlyGlnLysSerGlnThrProLeuSerProPro 340
DB 1021 GAACCTGCTCCTGCTGCTCAGAGGGTGAACAGAAAGAGAGAGAGCTTCACTCAGCC 1080
QY 341 CysAlaAspSerAlaThrSerGlyAlaAlaGlyLysArgProValHisProAlaSerVal 360
DB 1081 TGTGCCGACTCTCTCACTCAGAGGCGAGCAAGAGGCCCGTGCATCCCGCC----- 1124
QY 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnHisLeu 380
DB 1135 ---AGCTTACCCAGCGTGCAGCGCTCTTGAAGACAGCCGCTGGTACAGGCGCTC 1191
QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 400
DB 1192 AGTGGGCTGACCTGCTCGCAGAGAGCTGGAAGACAGCAATAGCTCAAGCGTTCTTCC 1251
QY 401 SerLeuAspHisLeuSerValSerTyrSerAlaSerMetAlaAspLeuHisGlyPhe 420
DB 1252 TCTCTGATATCAAAATCACTTATATTCAGCCAGCATGTCAGATCTTACATGCGCTTC 1311
QY 421 SerSerSerGlnAspAlaLeuGlyTyrTyrLysProSerThrHisLeuAspGlyHisAsn 440
DB 1312 TCTCTCATCAAGAAATGCTTTGAAATCTTCAAAACCTTCCACTCTGATGGGACCAAC 1371
QY 441 LysLeuCysGlnPheSerProValGlnGlyLeuSerGlnGlnThrProGlnThrSerPro 460
DB 1372 AAGGTATGCCAGTTCTCCCTGTTCAAGAACTATCGAGAGAGACTCCCGAAACAGTCTCT 1431

QY 461 AspLysGlyGlnAlaSerHisProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
DB 1432 GATTAAGAGAAAGCCAGCATTCCTCAAGAAAGCTGCAACCGCCAGGCTTTCAGACCCAG 1491
QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
DB 1492 AGCAAGCATTTGATTCGGTCAAGAACACAGACAGTGGACCGCCAGAGGTCCCTTTTA 1551
QY 501 SerProLeuHisArgSerGlySerValGlnAspAsnTyrHisThrSerPheLeuPheGly 520
DB 1552 TCTCCACTCATCGAAGTGGAGGCTGGAGAGAAATTAACCAACAGCTTCTTTTCGCG 1611
QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysLeuHisGly 540
DB 1612 CTTTCCACCAAGCCAGACGACCTTACGAAAGTCTGCTGGCTGGCTTAAAGGCTGGAC 1671
QY 541 SerAspHisLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrTyrPheAla 560
DB 1672 TGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTGATTTTGCC 1731
QY 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaHisTyrGlyLysSerAlaSerTyrSer 580
DB 1732 ACAGAGTCTCTCACTTCTTACTTCTGCTCAGCATCTTACGAGGAGAGCTGCACTACT 1791
QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
DB 1792 GCTTACAGCTGAGCCAGCTGCTGCTGCTGAGAACAAAGTCTTATCTGTTCCAGGCG 1851
QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrHisGlyLysSerProPheGly 620
DB 1852 CAGAACCAAGTACAGAGAGTGACTGCGCGGAGCTGGCATGAAGAGAGAGCCCTTTGAA 1911
QY 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyLysSerHisMetSerGlnAsn 640
DB 1912 AAGCACTTTTAAACGCAAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCAAGAAC 1971
QY 641 ArgSerArgGlnGlyLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGly 660
DB 1972 AGGTCAAGGAGAGAGCTGGGGAAGTGGGAGTCACTTACTTCTTCTGGGAGAGATGAA 2031
QY 661 HisLeuGlnValSer 665
DB 2032 ATCATTTAGAGTCTCC 2046

RESULT 14
US-10-072-012-255
; Sequence 255, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patnurajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Beha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Purtek, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsdorok II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same


```

FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072.012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO: 255
LENGTH: 2200
TYPE: DNA
ORGANISM: Homo sapiens
US-10-072-012-255

Alignment Scores:
Score: 0 Length: 2200
Pred. No.: 3358.50 Matches: 663
Best Similarity: 97.65% Conservative: 1
Best Local Similarity: 97.50% Mismatches: 1
Query Match: 98.26% Indels: 15
DB: 13 Gaps: 4

US-10-029-345a-109 (1-665) x US-10-072-012-255 (1-2200)
QY 1 MetAlahisglumetlleglthrglnlevalthrgluargleuvalaleuenglun 20
Db 61 ATGGCCCAAGATGATGAGTGAATCAATGTTACTGAGAGGTGGCTCTCTGGA 120
QY 21 SerGlyThrgluysvalleuenuleasSerArpPropheValGlyrYasrthSer 40
Db 121 AGTGAACGGAAGAGTGTCTGCTAATGATAGCCGCAATTTGTGAATCAATACATCC 180
QY 41 HislleuenglualaleanlleasnlySerlyseuWetlysaTgATgleuGln 60
Db 181 CACATTTTGAGACCAATTAATCAACTGCTCAAGCTTATGAAGCGAAGTTGCAACG 240
QY 61 AsplysvalleuilethrgluenuleGlnHisSerAlalyshslvs----- 76
Db 241 GACAAAGGTAAATTAACAGACTCATCCAGCATTCACGAAACATAAGTAAACGCTCAG 300
QY 77 ValAspIleasrCysserGlnlyseuValValTyraSpGlnSerSerGlnAspVala 96
Db 301 GTTGACATTTGATTCAGTCAAGAGTGTGTAGATTACGATCAAAAGCTCCCAAGATGTGC 360
QY 97 SerIleuSerArpCysserGlnlyseuValValTyraSpGlnlyseuSerPhean 116
Db 361 TCTTCTCTTCAGACTGTTTCTCACTGTTACTCTGGTAAACTGAGAAAGCTTCAAC 420
QY 117 SerValHisleuenule-----GlyGlyPheAlaGluPheSerArpGlyPhe 132
Db 421 TCTGTTCACTGCTTGAGGTTTATTCTTAAAGTGGGTTTGCGAGTCTCTCTGTTGTTT 480
QY 133 ProGlyLeuGlyGlnlyseuSerThrlleuValProthrCyseGlnProCyseu 152
Db 481 CTTGGCTCTGTGAAGAAATCACTCAAGTCCCTTACCTGATTTCTCACTTGTGCTTA 540

QY 153 ProValAlaAsnllglYProthrArgllleuProAsnleuTyrlleuGlyCyseGlnArg 172
Db 541 CTTGTTGCCAACAATTTGGCCCAACCCGAATTTCTCCCAATCTTAAATCTTGCTGCCAGCA 600
QY 173 AspValIleuAsnlyseuGlnleuileGlnGlnAsnGlylleGlyTyraValIleuAsnAlaSer 192
Db 601 GATGCTCTCAACAGAGACTGATGACAGAAATGGAATGTTATGTGTTAAATGCGACG 660
QY 193 TyrlthCyProlyPProAspPheIleProGlnSerHisPheleuArgValProValaen 212
Db 661 AATACCTGTCCAAAGCTGACTTATCTCCGAGCTTCAATTTCTGCGTGGCTGAGAT 720
QY 213 AspSerPheCyseGlnlyslleuProTrpIleuAspLyseuSerValAspPheIle----- 230
Db 721 GACAGCTTTGTGGAATAATTTGCGGTGTTGACAAATCAAGTATTTCTTGTAG 780
QY 231 -----GlyysAlaLysAlaSerAsnGlyCyseValleuValHisCyseleuAla 246
Db 781 TTGACTTATACAGAAAGCAAGCAAGCTCCAAATGATGTTCTAGTGCACATTTAGCT 840
QY 247 GlylleSerArgSerAlaThrIleAlaIleAlaTyrlleMetLysArgMetAspMetSer 266
Db 841 GGGATCTCCGCTCCGCCACCAATCGCTATCGCTTACATCAATGAAGATGACATGCT 900
QY 267 LeuAspGlnAlaTyr---ArgPheVallyseGlylysaTgPProthrlleSerProAsnPh 285
Db 901 TTAGATGAAGCTTACAGAGATTTGTGAAGAAAGAAACCTCACTATATCTCCAACTTC 960
QY 286 AsnPhelGlnGlyGlnleuLeuAspTyrllysaTgLyseuAsnGlnThrglyAlaSer 305
Db 961 AATTTTCTGGGCCCACTCTGACTATGAGAAAGATTAAGAACCAAGATCGAAGATCA 1020
QY 306 GlyProlyseuLyseuLyseuLeuenuleuGlnlyseuGlyProAsnGlnProValProAla 325
Db 1021 GGAGCCAAAGAGCAAACTCAAGCTGCTGACCTGAGAAAGCAATGAACCTGTGCTGCT 1080
QY 326 ValSerGlnGlyGlyGlnlyseuSerGlnThrProleuSerProProCyseAlaAspSerAla 345
Db 1081 GTCTCAAGAGGTGAGCAAGAAAGCAAGCCCTCAGTCACTGCTGCTGCTGCTGCT 1140
QY 346 ThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerValProSerValProSer 365
Db 1141 ACTTCAGAGGAGAGGAGCAAAAGCCCGTGCATCCGCCAGGCTGCCAGCTGCCAGC 1200
QY 366 ValGlnProSerleuLeuGlnleuAspSerProleuValGlnAlaLeuSerGlyLeuHisleu 385
Db 1201 GTGAGCCGTGCTGTTAGAGGACAGCCCGTGTACAGGCGCTCAGTGGCTGCACCTG 1260
QY 386 SerAlaAspArgleuGlnAspSerleuLyseuLyseuArgSerPheSerleuAspIleLys 405
Db 1261 TCCGACAGACGCTGGAAGACAGCAATTAAGCTCAAGCGTCTCTCTGATATCAAAA 1320
QY 406 SerValSerTyserAlaSerMetAlaAserleuHisGlyPheSerSerSerGlnAsp 425
Db 1321 TCGATTTCATATTCAGCCAGCATGTGCACATCTTACATGCTCTCTCATTCAGAAAT 1380
QY 426 AlaLeuGlnTyrlTyrllyseuProSerThrlleuAspGlyThrlleuLyseuGlnPhe 445
Db 1381 GCTTTGGAATATCAAAACCTTCACTTCTGATGAGGACCAACAAAGCTATGCCAGTTC 1440
QY 446 SerProValGlnGlnleuSerGlnGlnThrProGlnThrSerProAspLyseuGlnAla 465
Db 1441 TCCCTGTTCAAGAACTATCGGACACATCCCGAAACCAAGTCTGATTAAGAGAGAGCC 1500
QY 466 SerIleProLyseuGlnThrAlaArgProSerAspSerGlnSerlysaTgLeuHis 485
Db 1501 AGCATGCCCAAGAGGCTGCAGACCGAGGCTTCAAGACCGAGGCAAGGATGTCAT 1560
QY 486 SerValAlaGlnThrSerSerGlyThrlalGlnArgSerleuLeuSerProleuHisArg 505
Db 1561 TGGTTCAGAACCAACAGAGTGCACCGCCAGAGTCCCTTTTATCTCACTGATGCA 1620

```

QY 506 SerGlySerValGlnuPaentYrHsThSerPheLeuPheGlyLeuSerThSerGln 525
DB 1621 AGTGGAGCTGGAAGACATTTACACACCAAGCTTCTTTCCGCTTCCACAGCCAG 1680
QY 526 GlnHisLeuThrLysSerAlaGlyLeuGlyLysLeuGlyTyrPheLysSerAlaLeuAla 545
DB 1681 CACACCTCAAGAGCTGCTGGCTGGGCTTAAGGCTGGGACCTCGCATCTGGATCTTGCC 1740
QY 546 ProGlnThrSerThrProSerLeuThrSerSerTyrPheAlaThrGlnuSerSerHis 555
DB 1741 CCCACACCTCAAGCTGCTGGCTGGGCTTAAGGCTGGGACCTCGCATCTGGATCTTGCC 1800
QY 566 PheTyrSerAlaSerAlaLeuTyrGlyLysSerAlaSerTyrSerAlaTyrSerCysSer 585
DB 1801 TTCTACTCTGCTCAAGCATCTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
QY 586 GlnLeuProThrCysGlyLysSerGlnValTyrSerValArgArgArgGlnuLysProSerAsp 605
DB 1861 CAGCTGCCCACTTGGCGGAGCAAGCTCTATCTGTGGCGAGGCGGAGGAGGAGGAGGAGGAG 1920
QY 606 ArgAlaAspSerArgArgSerTyrPheLysGlnuSerProPheGlnuLysGlnuPheLysArg 625
DB 1921 AGAGCTGACCTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
QY 626 ArgSerCysGlnuMetGlnuPheGlyLysSerLysSerGlnuSerArgSerArgGlnu 645
DB 1981 AGAAGCTGACCAATTCGAATTTGGAGAGGAGCATCTGCAAGAGAGGAGGAGGAGGAGGAGGAG 2040
QY 646 LeuGlyLysValGlySerGlnuSerSerPheSerGlySerMetGlnuLeuLeuGlnuLysSer 665
DB 2041 CTGGGAGAGGTGGGAGGAGTCAAGTCTTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100

RESULT 15

US-09-964-277-20
Sequence 20, Application US/09964277
Patent No. US20020137170A1
GENERAL INFORMATION:
APPLICANT: Lucbe, Ralf M.
TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.434
CURRENT APPLICATION NUMBER: US/09/964,277
CURRENT FILING DATE: 2001-09-25
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 20
LENGTH: 3332
TYPE: DNA
ORGANISM: Homo sapiens
US-09-964-277-20

Alignment Scores:

Pred. No.: 8,866-317 Length: 3332
Score: 3048.50 Matches: 607
Percent Similarity: 91.44% Conservative: 2
Best Local Similarity: 91.14% Mismatches: 1
Query Match: 89.19% Indels: 56
DB: 9 Gaps: 1

US-10-029-345A-109 (1-665) x US-09-964-277-20 (1-3332)

QY 1 MetAlaHisGlnuMetLysGlyThrGlnuLeuAlaThrGlnuArgLeuValAlaLeuLeuGlnu 20
DB 562 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621
QY 21 SerGlyThrGlnuLysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 622 AGTGGAGCTGGAAGACATTTACACACCAAGCTTCTTTCCGCTTCCACAGCCAG 661
QY 41 HisLeuLeuGlnuAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB 682 CACATTTGGAGACCATTAATATCACTGCTCAAGGCTTATGAAGGAGGAGGAGGAGGAGGAG 741

QY 61 AspLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 742 GACCAAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 801
QY 81 CysSerGlnuLysValValValValValValValValValValValValValValValValVal 100
DB 802 TGCAGTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 861
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnuLysSerPheLeuSerValHisLeu 120
DB 862 GACTGTTTCTCACTGATCTTCTGGGTTAACTGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 921
QY 121 -LeuAlaGlyGlyPheAlaGlyPheSerArgCysPheProGlyLeuCysGlyLysSer 140
DB 922 CTGG----- 925
QY 140 ThrLeuValProThrCysLysSerGlnuProCysLeuProValAlaAlaMetLysGlyProth 160
DB 925 ----- 925
QY 160 ArgGlyLeuProAlaMetLysGlyCysGlnuArgPheValLeuAlaMetLysGlyLeuLeu 180
DB 926 -----CAGGAGCTGAT 936
QY 180 eGlnGlnuMetLysGlyTyrValLeuAlaMetLysSerTyrThrCysProLysProAspPhe 200
DB 937 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 996
QY 200 eLysProGlnuSerHisPheLeuArgValProValAlaMetSerPheCysGlnuLysLeu 220
DB 997 TATCCCGAGTCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
QY 220 UP-OT-PL-LEU-AP-LEU-AP-LEU-AP-LEU-AP-LEU-AP-LEU-AP-LEU-AP-LEU-AP-LEU 240
DB 1057 GCCGTGTTGGAACCAATTCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1116
QY 240 LeuValHisCysLeuAlaGlyLysSerArgSerAlaThrLeuAlaLeuAlaLeuAlaLeuAla 260
DB 1117 TCTAGTCACTGTTTACGCTGGATCTCCGCTCCGACCATGCTGCTGCTGCTGCTGCTGCTGCT 1176
QY 260 LysArgMetAspMetSerLeuAspGlnuAlaTyrArgPheValLysGlnuLysArgProth 280
DB 1177 GAAGAGATGACATCTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1236
QY 280 LysSerProAlaMetPheLeuGlyLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 300
DB 1237 TATATCTCCAACTTCATTTCTGGGCCCAACTCTGACATGAGAGAGAGATTAAGAA 1296
QY 300 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlnuLysProAsp 320
DB 1297 CCAAGCTGAGCATCAAGGCCCAAGGCAAACTCAAGCTGCTGACCTGAGAGAGCCAA 1356
QY 320 GlnuProValProAlaValSerGlnuGlyLysGlnuLysSerGlnuThrProLeuSerProPhe 340
DB 1357 TGAACCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1416
QY 340 OCYAlaAspSerAlaThrSerGlnuAlaAlaGlyLysGlnuArgProValHisProAlaSerVal 360
DB 1417 CTGTGCCGACTCTGCTCACTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1476
QY 360 LysProValProSerValGlnuProSerLeuGlnuAspSerProLeuValGlnuAlaLeu 380
DB 1477 GCCCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1536
QY 380 LysGlyLeuHisLeuSerAlaAspArgLeuGlnuAspSerAlaLysLeuLysArgSerPhe 400
DB 1537 CAGTGGCTGACACTGTCGAGAGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1596
QY 400 eSerLeuAspLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
DB 1597 CTCTCGATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1656
QY 420 eSerSerSerGlnuAspAlaLeuGlnuTyrTyrLysProSerThrThrLeuAspGlyThrAsp 440

Db 1657 CTCCTCATCAGAGATGCTTTGGAAATACAAACCTTCCACTCTGATGGAGCCAA 1716
Qy 440 nlyslleucysglnpheserprovalglnluleuserglnlnrproglunthrserpr 460
Db 1717 CAAGCTATGCAAGTTCCTCCCTGTTCAGAGAACTATCGAGAGACTCCCGAAACCAAGTCC 1776
Qy 460 oasplysglnlualaserlieprolyslleuqlnlnrlnlaargproseraspsergl 480
Db 1777 TGATTAAGAGAGAAAGCCAGATCCCAAGAGCTGCAAGCCGACGCTTCAGACAGCA 1836
Qy 480 nserlysalgleuhservalarglnrsersergilythralaglnargserleue 500
Db 1837 GAGCAGAGCATTCGATTCGATCAGAACAGCAGAGAGTGGACCCGCCAGAGTCCCTTTT 1896
Qy 500 userproleuhsargserglservalgluaspsantyrhlsrserphleuphegl 520
Db 1897 ATCTCCACTGCATCCAAAGTGGAGGCTGAGAGCAATTAACACACAGCTTCCCTTTCCG 1956
Qy 520 yleuserthrserglnlnhsleuthrlysseralaglyleuqlyleuylsglytrphl 540
Db 1957 CTTTCCACAGCCAGCAGCACTCACGAAGTCTGCTGGCTGGCTTAAAGGCTGGCA 2016
Qy 540 sseraspilseuualaproglunthrserthrproserleuthrsersertrpyrphal 560
Db 2017 CTCGGATATCTTGCCCCCAGACCTCTACCCCTTCCTGACACAGCAGCTGGATTTTGC 2076
Qy 560 athrghluserserhispherlyseralaserlailetyrghlygllyseralaserlyse 580
Db 2077 CACAGAGTCTCACACTTCTACTCTGCTCAGCCATCTAGAGAGCAGTGCAGTTACTC 2136
Qy 580 rlaalyrsercysserglnleuprophrlysglyaspglnvaltyrservalargargar 600
Db 2137 TGCCTACAGCTGACGCGAGCTGCCCACTTGCGGAGACCAAGTCTATTCTGTGCGCAGGCG 2196
Qy 600 gglunlypserasparghlaspserargysetrphlsglunluserprophel 620
Db 2197 GCAAGAGCCAAAGTACAGAGCTGACTCGCGCGAGCTGCATGAAGAAGGCCCTTTGA 2256
Qy 620 ulysglnphelyargargsercyglnmetglnpheglvglyuserllemetsergluas 640
Db 2257 AAAGCAGTTTAAACGACGAAGCTGCCAATGGAATTTGAGAGAGCATCATGTCAAGAA 2316
Qy 640 naqserarglnluleuqllysvaiglyserglnserserpheserghlysermetgl 660
Db 2317 CAGGTACCGGAAAGCTGGGAAAGTGGGAGTCAGTCTAGCTTTTGGGACAGCATGGA 2376
Qy 660 uileileglnvalser 665
Db 2377 AATCANTGAGTCTCC 2392

Search completed: June 22, 2004, 03:06:35
Job time : 857.731 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 19:56:46 ; Search time 735.833 Seconds
(without alignments)
3839.259 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 3418
Sequence: 1 MAHEMTCGTQVTERVALLE.....LGKVGSGSRFSGSMREIIVS 665

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 segs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-Q/cgnt2.1/USPTO.spool/US10029345/runat.21062004.164145.29020/app_query.fasta_1.1294
-DB=Geneseg.29Jan04 -OPMT=fastac -SUFFIX=ring -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345.QCGN.1.1.687/runat.21062004.164145.29020 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseg.29Jan04:*

1: genesegq1980s:*\n2: genesegq1990s:*\n3: genesegq2000s:*\n4: genesegq2001as:*\n5: genesegq2001bs:*\n6: genesegq2002as:*\n7: genesegq2003as:*\n8: genesegq2003cs:*\n9: genesegq2003cs:*\n10: genesegq2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3418	100.0	5450	6	ACC60559 Polynucle
2	3418	100.0	5450	6	ACC60572 Polynucle
3	3406	99.6	2732	4	AAD09492 Human SGP
4	3406	99.6	3059	6	AAS15768 CDNA seg
5	3406	99.6	3104	6	ABN59704 Novel hum
6	3406	99.6	3496	6	ABK47596 CDNA enco
7	3406	99.6	3544	5	AAS14639 Human CDN
8	3406	99.6	3544	6	ABK49402 CDNA enco

153-277 Check

9	3406	99.6	3766	6	ABK14474	ABK14474 Human pro
10	3406	99.6	4790	6	ABN83966	ABN83966 Human gen
11	3406	99.6	5145	5	ABV20833	ABV20833 Human pro
12	3406	99.6	5145	5	ABV21080	ABV21080 Human pro
13	3406	99.6	5145	5	ABV26680	ABV26680 Human pro
14	3406	99.6	5145	5	ABV20978	ABV20978 Human pro
15	3406	99.6	5145	5	ABV21092	ABV21092 Human pro
16	3406	99.6	5145	5	ABV21312	ABV21312 Human pro
17	3406	99.6	5145	5	ABV21316	ABV21316 Human pro
18	3406	99.6	5145	5	ABV26826	ABV26826 Human pro
19	3406	99.6	5145	5	ABV27131	ABV27131 Human pro
20	3406	99.6	5145	5	ABV26923	ABV26923 Human pro
21	3406	99.6	5145	5	ABV27135	ABV27135 Human pro
22	3406	99.6	5145	5	ABV28657	ABV28657 Human pro
23	3406	99.6	5145	5	ABV22827	ABV22827 Human pro
24	3406	99.6	5145	5	ABV26934	ABV26934 Human pro
25	3399	99.4	2102	7	ADAS3105	ADAS3105 Human cod
26	3398	99.4	2966	4	AAH93685	AAH93685 Human pro
27	3388	99.1	5111	6	ACC60521	ACC60521 Polynucle
28	3373	98.7	2118	4	AAFP0479	AAFP0479 Human pro
29	3057.5	89.5	2756	6	ACC60560	ACC60560 Polynucle
30	3048.5	89.2	3332	6	ABK48378	ABK48378 CDNA enco
31	1326	38.8	2377	7	ACA64956	ACA64956 Human pro
32	1326	38.8	2377	7	ABX10760	ABX10760 Human dua
33	1302	38.1	2453	6	AAH41236	AAH41236 Murine ne
34	1297	37.9	2453	2	AAT86758	AAT86758 CDNA of t
35	1262	36.9	2476	4	AAS31013	AAS31013 Human dia
36	1246	36.5	749	4	AAH06539	AAH06539 Human cod
37	1164.5	34.1	2415	2	AAT86757	AAT86757 CDNA of t
38	977.5	28.6	1917	9	ADDE0745	ADDE0745 Novel cod
39	742.5	21.7	1171	5	AAS72089	AAS72089 DNA encod
40	724	21.2	418	5	ABV11252	ABV11252 Human pro
41	724	21.2	461	5	ABV40981	ABV40981 Human pro
42	724	21.2	461	5	ABV42398	ABV42398 Human pro
43	724	21.2	461	5	ABV41327	ABV41327 Human pro
44	707	20.7	425	5	ABV10726	ABV10726 Human pro
45	703	20.6	467	5	ABV40998	ABV40998 Human pro

ALIGNMENTS

RESULT 1	ACCG60559	standard; CDNA; 5450 BP.
ID	ACCG60559	
AC	ACCG60559;	
DT	19-JUN-2003	(first entry)
XX		
DE	Polynucleotide relating to the invention SEQ ID NO: 108.	
XX		
KW	Gene; ss; antiproliferative; hepatotropic; nephrotropic; antitachytic; antiproliferative; cardiatic; cytosolic; gene therapy; liver disease; immunological disorder; arthritis; psoriasis; congenital heart defect; congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200257460-A2.	
XX		
PD	25-JUL-2002.	
XX		
PF	20-DEC-2001; 2001WO-US050459.	
XX		
PR	20-DEC-2000; 2000US-0256868P.	
PR	30-MAR-2001; 2001US-0280186P.	
PR	01-MAY-2001; 2001US-0287735P.	
PR	05-JUN-2001; 2001US-0295848P.	
PR	25-JUN-2001; 2001US-0300465P.	
XX		
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.	
XX		


```

Db      2098 CTTTCACAGCAGCAGCAGCAGCTCAGCAAGTCTGTGCTGGGCTTAAGGCTGGCAC 2157
Qy      541 SerAaplleuAlaProGlnThrSerThrProSerLeuThrSerSerThrPyrPheAla 560
Db      2158 TCGGATATCTGGCCCCCAGACCTTACCCCTCCCTGACAGCAGCTGGATATTTTGCC 2217
Qy      561 ThrGlnSerSerHisPheThrSerAlaSerAlaIleTyrGlySerAlaSerTyrSer 580
Db      2218 AAGAGCTCTTCACTTCTTACTCTGCTTACCCATCTTACGAGGACAGTCCAGTTACTCT 2277
Qy      581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
Db      2278 GCCTACAGCTGACAGCAGCTGCCACTTGCGGAGCAACAGTGTATCTGTGGCAGGCGG 2337
Qy      601 GlnTyrProSerAspArgAlaAspSerArgArgSerThrHisGlnGlnSerProPheGln 620
Db      2338 CAGAAAGCCAGATGACAGACCTGACTCGCGCGAGCTGACATGAAAGAGCCCTTTGAA 2397
Qy      621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAan 640
Db      2398 AAGCAGTTTAAAGCAGAAAGCTGCCAAATGAAATTTGAGAGAGCATGATGACAGAAC 2457
Qy      641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
Db      2458 AGGTACCGGGAAGAGCTGGGGAAGTGGGCAATGACTTACGTTTTCGGGCAAGCATGAA 2517
Qy      661 IlelleGlnValSer 665
Db      2518 ATCATTGAGCTCTCC 2532

RESULT 2
ACC60572
ID      ACC60572 standard; cDNA, 5450 BP.
XX      AC
XX      ACC60572;
DT      19-JUN-2003 (first entry)
XX      DE
XX      Polynucleotide relating to the invention SEQ ID NO: 147.
XX      KW
XX      Gene; sb; antiproliferative; hepatocytic; nephrotropic; antiarthritic;
XX      antiproliferative; cardiatic; cystostatic; gene therapy; liver disease;
XX      immunological disorder; arthritic; psoriasis; congenital heart defect;
XX      congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX      OS
XX      Homo sapiens.
XX      PN
XX      WO200257460-A2.
XX      PD
XX      25-JUL-2002.
XX      PF
XX      20-DEC-2001; 2001WO-US050459.
XX      PR
XX      20-DEC-2000; 2000US-0256868P.
XX      PR
XX      30-MAR-2001; 2001US-0280186P.
XX      PR
XX      01-MAY-2001; 2001US-0287735P.
XX      PR
XX      05-JUN-2001; 2001US-0295848P.
XX      PR
XX      25-JUN-2001; 2001US-0300465P.
XX      PA
XX      (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX      PI
XX      Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
XX      PI
XX      Siemere N, Bol D, Schieven G, Finger J, Todderud CG, Basolino D,
XX      PI
XX      Kryetsek S, Mcatee P, Suchard S, Banas D;
XX      DR
XX      MPI: 2002-599721/64.
XX      P-PSDB: ABR52407.
XX      PT
XX      Novel polynucleotides encoding human phosphatase polypeptides useful in
XX      the prevention or treatment of e.g. proliferative and cardiovascular
XX      disorders.

```

```

PS      Example 7; Fig 19; 801bp; English.
XX      CC
XX      The invention relates to a novel isolated nucleic acid comprising a
XX      polynucleotide having a nucleotide sequence selected from 40
XX      polynucleotides fully defined in the specification. The polynucleotide of
XX      the invention has antiproliferative, hepatocytic, nephrotropic,
XX      antiarthritic, antipsoriatic, cardiatic, and cystostatic activity. The
XX      polynucleotide may have a use in gene therapy. A polynucleotide or
XX      polypeptide of the invention is useful for preventing, treating or
XX      ameliorating a medical condition, e.g. a proliferative disorder. They are
XX      also useful for treating e.g. liver disease, renal failure, immunological
XX      disorders including arthritis and psoriasis, cardiovascular disorders
XX      such as congenital heart defects and congestive heart failure, and
XX      cancer. A method of the invention is useful for diagnosing a pathological
XX      condition or susceptibility to a condition in a subject. The present
XX      sequence is used in the exemplification of the invention.
XX      SQ
XX      Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;
XX      Alignment Scores:
XX      Pred. No.: 3,02e-235 Length: 5450
XX      Score: 3418.00 Matches: 665
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 6 Gaps: 0
XX      US-10-029-345A-109 (1-665) x ACC60572 (1-5450)
Qy      1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
Db      538 ATGCCCATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 597
Qy      21 SerGlyThrGlnLysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db      598 AGTGAACAGGAAAGAAAGTGGCTTAATTGATGATGATGATGATGATGATGATGATGATGAT 657
Qy      41 HisIleLeuGlnAlaIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
Db      658 CACATTTTGGAAAGCCATTAATATATCACTGCTCAAGCTTATAGAGGAGGATGGCAACAG 717
Qy      61 AapLysValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisLysValAspIleAsp 80
Db      718 GACAAAGCTTATATACAGAGCTATCCAGATTACGAGAAACATAGATGATGATGATGAT 777
Qy      81 CysSerGlnLysValValValIleValIleValIleValIleValIleValIleValIleVal 100
Db      778 TGCAGTCAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
Qy      101 AapCysPheLeuThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
Db      838 GACTGTTTCTCACTGATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 897
Qy      121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLysGlyLysSer 140
Db      898 CTTCAGAGTGGGTTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 957
Qy      141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db      958 ACTCTAGTCCCTACCTGATCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1017
Qy      161 ArgIleLeuProAsnLeuThrLeuGlyCysGlnArgAspValLeuAsnIleGlyLeuIle 180
Db      1018 CCAATTTCTCCCAATCTTATCTTGGCTGCAGCAGAGATGCTCTCAACAGAGAGCTGAATA 1077
Qy      181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
Db      1078 CAGCAGAAATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137
Qy      201 IleProLysSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
Db      1138 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGAGAAATTTTG 1197

```


QY 221 ProTribLeuAspLysSerValaasphe11egluLysAlaLysAsnGlyCysVal 240
 Db 1198 CCGGTGTTGGAACAAATGATGATTTCTATGAGAAAGAAAGCCCTCCAAATGATGCTTT 1257
 QY 241 LeuValH1sCysLeuAlaGly11eSerArgSerAlaThr11eAla11eAla11eMet 260
 Db 1258 CTAGTGCATGTTTACTGCTGGATCTCCCGCTCCGACCAATGCTATGCGCTCAATCATG 1317
 QY 261 LysArgMetAspMetSerLeuaspGlu1a1YrArgPheValLysGluLysArgProThr 280
 Db 1318 AAGAGATGAGCATGCTTTAGATGAGCTTACATATTGTGAAAGAAAGAAAGCTTACT 1377
 QY 281 11eSerProAsnPhenAsnPhenLeuGly11eLeuAsp1YrGlyLys11eLysAsn 300
 Db 1378 AATATCTCCAACTTCAATTTCTGGCCAACTCTCGACTTATGAGAAAGATTAAGAAC 1437
 QY 301 GlnThrGlyAlaSerGly1ProLysSerLysLeuLysLeuH1sLeuGlyLysProAsn 320
 Db 1438 CAGACTGAGCATGCTCAGGGCCAAAGAGCAAACTCAAGCTGCTGCATCGAGAAAGCCAAAT 1497
 QY 321 GluProVal1ProAlaValaSerGlyGly1YrGlyLysSerGlyLysProLeuSerProPro 340
 Db 1498 GAACCTGTCCTGCTGCTCAGAGGGTGGACAGAAAGCAAGCCCTCACTTCCACCC 1557
 QY 341 CysAlaAspSerAlaThrSerGlyAlaAlaGly1YrArgProValH1sProAlaSerVal 360
 Db 1558 TGTGCCACTGCTGCTCACTCAGAGGCAAGAGCAAAAGGCCGTCATCCCGCAGCCGTG 1617
 QY 361 ProSerVal1ProSerVal1GlnProSerLeuLeuGly1AspSerProLeuVal1GlnAlaLeu 380
 Db 1618 CCCAGCGTGGCCAGCGTGCAGCGCGTCTTGAAGAGCAAGCCCGTGTATCAGAGCGCTC 1677
 QY 381 SerGlyLeuH1sLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
 Db 1678 AGTGGGCTGCACTGTGTCGAGCAGAGCTGGAAGAGCAATATGCTCAAGGTTCTTCC 1737
 QY 401 SerLeuAsp11eLysSerValaSer1YrSerAlaSerMetAla11eSerLeuH1sGlyPhe 420
 Db 1738 TCTCTGATATCAATCACTTTCATATTCAGACAGATGCGAGCATCTTTCATGAGGCTTC 1797
 QY 421 SerSerSerGlyAspAlaLeuGlyLys1YrLysProSerThrThrLeuAspGly1YrAsn 440
 Db 1798 TCTCTATCAGAAAGATGCTTGGAAATATCAAAACCTTCACTCTGAGTGGAGCCAAAC 1857
 QY 441 LysLeuCysGlnPheSerProVal1GlnLysLeuSerGlyLysThrProGlnThrSerPro 460
 Db 1858 AAGCTATGCGAGTTCCTCCCTGTTCAAGAACTATCGAGAGACTCCCGAAACCAAGTCTT 1917
 QY 461 AspLysGluGluAlaSer11eProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 Db 1918 GATTAGAGAGAAAGCCAGATCCCAAGAGCTGCAAGACCCCGAGGCTTCAAGACACCAAG 1977
 QY 481 SerLysArgLeuH1sSerValaArgThrSerSerSerGly1YrAlaGlnArgSerLeuLeu 500
 Db 1978 AGCAAGCAGTATGCTTGGTCAAGAACAGAGCAAGTGGACCCCGCCAGAGGTCCTTTTAA 2037
 QY 501 SerProLeuH1sAspSerGlySerVal1GluAspAsn1YrH1sThrSerPheLeuPheGly 520
 Db 2038 TCTCACTGCAATCGAATGTGAGGTGGAGAGCAATTAACAACAAGCTTCTTTTCCGC 2097
 QY 521 LeuSerThrSerGlnGlnH1sLeuThrLysSerAlaGlyLeuGly1YrLysGly1YrH1s 540
 Db 2098 CTTTCCACAGCCAGCAAGCACTCAAGAGTCTGCTGGCTTAAAGGCTGGAGC 2157
 QY 541 SerAsp11eLeuAlaProGlnThrSerThrProSerLeuThrSerSer1YrPheAla 560
 Db 2158 TCGGATATCTTGGCCCCCAGAGCTTACCCCTCCCTGACAGAGCTGATATTTTGGCC 2217
 QY 561 ThrGluSerSerH1sPhe1YrSerAlaSerAla11e1YrGly1YrSerAlaSer1YrSer 580
 Db 2218 ACAAGTCTCTCAACTTACTCTGCTCAAGCACTTCAAGAGGCAAGGCTTACTCT 2277
 QY 581 Ala1YrSerCysSerGlnLeuProThrCysGlyAspGlnVal1YrSerValaArgArg 600

Db 2278 GCCTACAGCTGACGACGCTGCCACTTCCGGAACCAAGTCTATTTCTGTGCGCAGCGG 2337
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerThrPheGluGlyLysProPheGly 620
 Db 2338 CAGAAAGCCAAAGTACAGAGCTGACTCGCGGAGCTGGCATGAAAGAGCCCTTTGAA 2397
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGlyPheGly1YrGlyLysSerGlyLys 640
 Db 2398 AAGCGTTTAAACGCAAGAGCTGCCAAATGATTTGAGAGAGCATATGTCAGAGAAC 2457
 QY 641 ArgSerArgGluGluLysGlyValaGlySerGlnSerSerPheSerGlySerMetGly 660
 Db 2458 AGCTCAAGGAGAGCTGGGAGAAAGTGGGAGTCACTGACTTTCGGGCAAGCATGGA 2517
 QY 661 11eLeuGluValSer 665
 Db 2518 ATCATGAGGCTCTCC 2532
 RESULT 3
 AAD09492
 ID AAD09492 standard; DNA; 2732 BP.
 AC AAD09492;
 XX
 DT 10-SEP-2001 (first entry)
 XX
 DE Human SGP002 phosphatase polypeptide encoding DNA.
 XX
 KW Human; SGP002 phosphatase polypeptide; phosphatase-related disease;
 KW immune-related disorder; ocular disease; organ transplant rejection;
 KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
 KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
 KW attention disorder; cognition disorder; psychotic disorder; cytoskeletal;
 KW neurological disorder; vitreous; nocturnal; cerebroprotective therapy;
 KW neuroprotective; antibacterial; vulvar; tranquilizer; antispasmodic;
 KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
 KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
 KW MKP; migraine; chromosome 12p11.1-p12.1; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 538..2535
 FT /*tag= a
 FT /product= "Human SGP002 phosphatase polypeptide"
 XX
 PN W0200146394-A2.
 XX
 PD 28-JUN-2001.
 XX
 PP 21-DEC-2000; 2000MO-US034736.
 XX
 PR 21-DEC-1999; 99US-017325SP. NO
 PR 28-DEC-1999; 99US-0175766P. NO
 PR 25-JAN-2000; 2000US-0178078P. NO
 PR 31-JAN-2000; 2000US-0179301P. NO
 XX
 PA (SUGR-) SUGEN INC.
 XX
 PI P Bowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;
 PI Hill RJ, Flanagan P;
 DR MPI; 2001-418058/44.
 DR P-PSDB; AAE04834.
 XX
 PT Novel phosphatase polypeptide useful for treating cancers, immune-related
 PT diseases and disorders, cardiovascular disease, brain or neuronal-
 PT associated diseases and metabolic disorders.
 XX
 PS Claim 29; Fig 1; 166pp; English.

XX The present invention relates to phosphatase polypeptides, nucleotide sequences encoding them, as well as various products and methods useful for the diagnosis and treatment of various phosphatase-related diseases and conditions. Substance that modulates the activity of phosphatase polypeptide is used to treat immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases and metabolic disorders, including cancers of tissues, cancers of haematopoietic origin, diseases of central and peripheral nervous system, Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, viral infections, infections caused by prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual dysfunction, mood disorders, attention disorders, cognitive disorders, hypotension, hypertension, psychotic disorders, neurological disorders, dyskinesias and organ transplant rejection. The present sequence is a DNA encoding human SGP002 phosphatase polypeptide. This sequence is classified as dual specificity phosphatase (DSP) and MAP kinase phosphatase (MKP). SGP002 gene maps to chromosomal position 12p11.1-p12.1

XX Sequence 2732 BP, 710 A; 684 C; 686 G; 652 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8,73e-235 Length: 2732
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x AAD09492 (1-2732)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
DB 538 ATGGCCCATGAGATGATGGACATCAATGTTAATGAGAGTGGGCTGCTGCGAA 597
QY 21 SerG1ThrcGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
DB 598 AGTGGACGAAAGAGCTGCTTAATGATAGCGGCATTTGGAAATACATACATCC 657
QY 41 HisIleGluGlnAlaIleAsnIleAsnGlySerIleLeuMetIleArgArgLeuGln 60
DB 658 CACATTTGGACCATTAATATCACTGCTCCAAAGCTTATGAAGCAAGGTTGCAACAG 717
QY 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleValIleValIleAsp 80
DB 718 GACAAAGGTGTAATATACAGAGCTATCCACATTCACGAAACATAGGTGACATTGAT 777
QY 81 CysSerGlnIleValValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 778 TGCAGTCAGAAAGTGTGATTTACGATCAAGCTCCCAAGATGTTCCTCTCTTCA 837
QY 101 AspCysPheLeuThrValLeuLeuGlyIleValLeuGlnIleValSerPheAsnSerValHisLeu 120
DB 838 GACTGTTTCTCACTGATCTTGGGTAACCTGGAGAAAGACTTCAACTGTTCACTG 897
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyIleValSer 140
DB 898 CTTCAGAGTGGGTTGCTGATGCTCTCTGTTTCCCTGGCCCTCTGTGAAGAAATCC 957
QY 141 ThrLeuValProThrCysHisSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 958 ACTCTAGTCCCTACCTGATTTCTCAAGCTTGGCTTACTTGGCAACATTTGGCAAC 1017
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGlyLeuIle 180
DB 1018 CGAATCTTCCCATCTTATCTTGGCTGCAAGGATGATCCCAACAGAGAGCTGATG 1077
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProIleProAspPhe 200
DB 1078 CACGAGATGGGATGGTGTATGTTAAATGCGAGCAATACCTGTCCAAAGCCGTGATTT 1137
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleLeu 220

DB 1138 ATCCCGAGTCTCATTTCTGCTGCTGCTGATATGACAGACTTTTGTGAGAAATTTTG 1197
QY 221 ProTTPLeuAspIleSerValAlaPheIleGlnIleValAlaIleValAlaSerAsnGlyCysVal 240
DB 1198 CCGTGTGGGCAAAATCAGTATGATTTCTATGAAAGCAAAAGCCCTCAATGATGGTGT 1257
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleValIleMet 260
DB 1258 CTATGATCTGTTTACTGGGATCTCCGCTCCGCAACATCTGCTATGCTTACTATG 1317
QY 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValIleGlyIleValArgProThr 280
DB 1318 AAGAGATGAGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 1377
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnIleLeuAspTyrGluIleValIleValAsn 300
DB 1378 ATATCTCCAAACTTCAATTTTCTGGGCACTCTGACATATGAAAGAAATTAAGAAC 1437
QY 301 GlnThrGlyAlaSerGlyProIleSerIleValLeuIleLeuHisLeuGlnIleValProAsn 320
DB 1438 CAGACTGAGACATCAGGCGCAAGCAACCACTAAGCTGCTGCACTGGAGAAAGCCAAAT 1497
QY 321 GluProValProAlaValSerGluGlyGlnIleValSerGluThrProLeuSerProPro 340
DB 1498 GAACCTGCTCCTGCTGCTCAGAGGGTGGACAGAAAAAGCAGACGCCCTCACTCCACC 1557
QY 341 CysAlaAspSerAlaThrSerGlnAlaIleGlyGlnArgProValHisProAlaSerVal 360
DB 1558 TGTGCGACTGCTGCTACCTCAGAGGAGCAGAGCAAAAGCCCGTGCATCCCGCAGGTG 1617
QY 361 ProSerValProSerValGlnProSerLeuGlnAspSerProLeuValGlnAlaLeu 380
DB 1618 CCGAGGCTGCCAGCGGAGCGCTGCTGTATGAGACAGCCCGTGTACAGCGCTC 1677
QY 381 SerG1LeuHisLeuSerAlaAspArgLeuGluAspSerAsnIleValIleValIleValSerPhe 400
DB 1678 AGTGGCTGACCTGCTGCTGCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCGTCTTC 1737
QY 401 SerLeuAspIleIleValSerValSerTyrSerAlaSerMetAlaIleSerLeuHisGlyPhe 420
DB 1738 TCTCTGATATCAAAATCAGTTTCAATTCATTCAGCAGACATGACATCTTACATGCTTC 1797
QY 421 SerSerSerGlnAspAlaLeuGluTyrTyrIleProSerThrThrLeuAspGlyThrAsn 440
DB 1798 TCTCTATCAGAAATGCTTTGAAATCTACAACTTCCATCTCTGATGGAGCAAC 1857
QY 441 LysLeuCysGlnPheSerProValGlnIleValSerGlnIleThrProGluThrSerPro 460
DB 1858 AAGCTATGCAAGTCTCCCTGTTCAAGAACTATCGAGACAGCTCCGAAACAGTCT 1917
QY 461 AspIleGluGlnAlaSerIleProIleValLeuGlnIleThrAlaArgProSerAspSerGln 480
DB 1918 GATTAAGAGAGAACCCAGATCCCAAGAGCTCAGACTGCGGCTTACAGACAGCAG 1977
QY 481 SerIleArgLeuHisSerValArgThrSerSerSerG1ThrAlaGlnArgSerLeuLeu 500
DB 1978 AGCAAGAGATGATTCGGTCAAGAACCAAGAGAGAGGACCCGCCAGAGATCCCTTTA 2037
QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
DB 2038 TCTCCATGATGAGATGGAGGAGCGTGAAGCAATTAACACACAGCTTCCTTTTGGC 2097
QY 521 LeuSerThrSerGlnGlnHisLeuThrIleValSerAlaGlyIleValIleValIleValIle 540
DB 2098 CTTCACACAGCAGACAGCCTCAGAAAGTCTGCTGCGCTGAGCTTAAAGGCTGAC 2157
QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrIlePheAla 560
DB 2158 TCGGATATCTGGCCCCCAGACCTTACCTTCCCTGACACAGACGATGATTTTGGC 2217
QY 561 ThrGluSerSerHisPheTyrSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
DB 2218 ACAGAGTCCCAACTTACTGCTGCTGAGCATTAAGGAGGAGTGCAGTACTCT 2277

Oy	581	AlaIySerCysserGluLeuP ^o hm ^o YsgIyaSpGlnVal ^l TySerVala ^r ArgArg ^r	600
Db	2278	GCCrKAGcctGcAGcCGcCTGCCcATTCCGAGACCAAGCTATTCTGTGGCCAGCGG	233
Oy	601	GlnIyProSerA ^r pa ^r rg ^l Ala ^r SpSer ^r a ^r g ^r Arg ^r Se ^r T ^r ph ^l sgI ^l glnI ^l uSer ^r Pro ^l ph ^l gln	620
Db	2338	CAGAAgCCAAGTACAGAGCTGACTCGCGCGAGACTGGCAtGAAAGAGACCCCTTTGAA	2397
Oy	621	Ly ^g lnPhe ^l ya ^r rg ^r Arg ^r Ser ^r Cy ^g sl ^l met ^l glnPhe ^l gIy ^l uSer ^l l ^l me ^r Se ^r gI ^l ua ⁿ	640
Db	2398	AAAGAGTTTAAAGCGCAAGAGCTGCCAAATGGAAATTTGAGAGAGCATCAATGTCAGAGAAC	2455
Oy	641	ArgSer ^r ArgGlnI ^l uLeu ^l gI ^l ysVal ^l gI ^l Ser ^l glnSer ^r ph ^l Se ^r gI ^l ys ^r er ^l net ^l gln	660
Db	2458	AGGTACCGGGAAGAGCTGGGGAAGTGGCAGTCA ^r GTCTTTCGGGAGCAATGAA	2517
Oy	661	l ^l leI ^l leGlnVal ^l Ser	665
Db	2518	ATCATTTAGAGTCTCC	2532

DR WPI; 2002-010917/01.
P-PsDB; AAU09946.

XX Novel dual specificity phosphatase polypeptides useful for treating
PT cancer, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
PT disease, depression, schizophrenia, asthma and immune disorders.

XX Claim 5; Page 34-37; 43pp; English.

CC The present invention relates to a new isolated dual specificity
CC phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
CC sequence that is fully defined in the specification. The invention also
CC provides a sequence encoded by a 3059 nucleotide sequence fully defined
CC in the specification, and a sequence having at least 95 % identity to the
CC polypeptide, or fragments or variants of DUSP-10. The invention is useful
CC for treating cancer e.g. leukaemia, colon carcinoma, lung cancer,
CC prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
CC stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,

CC schizophrenia, cardiac myotrophias, asthma, immune disorders,
CC inflammatory processes e.g. arthritis, bowel diseases, type I diabetes,
CC osteoporosis, diabetes and diabetes associated diseases. The molecules o
CC the invention are also useful as vaccines for inducing immunological
CC response in a mammal, in disease diagnosis and in assays for screening
CC agonistic or antagonistic compounds. Other uses of the invention include
CC identifying membrane bound or soluble receptors, as a diagnostic reagent
CC in chromosome localisation studies, and as a valuable tool in tissue
CC expression studies. The present sequence represents cDNA of the human
CC dual specificity phosphatase, DUSP-10. DUSP-10 is located on chromosome
XX 12

sq Sequence 3059 BP; 831 A; 761 C; 709 G; 758 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.02e-234	Length:	3059
Score:	3406.00	Matches:	663
Percent Similarity:	99.85%	Conservative:	1
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	99.65%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345A-109 (1-665) x AAAS15768 (1-3059)

```

Db      847 CTATGACACTGTTAGTGGGATCTCCGCTCCGCCACCATCGCTACATCATG 906
Qy      261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleuSerGluAsn 280
Db      907 AAGAGATGAGCATGCTCTTAGATGAAGACTTACAGATTTGTGAAGAAAAAAGACCTACT 966
Qy      281 IleSerProAsnPheAsnPheLeuGluIleuLeuAspTyrGluValIleuAsn 300
Db      967 ATATCTCCAACTCAATTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAGAAC 1026
Qy      301 GluThrGluAlaSerGlyProLysSerLeuIleuLeuAsnIleuGluIleuProAsn 320
Db      1027 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAGCTGCTGCACTCGGAGAAACCAAT 1086
Qy      321 GluProValProAlaValSerGluGlyIleuLysSerGluThrProLeuSerProPro 340
Db      1087 GAACGTGTCCCTGCTGTCTGAGAGGTGACAGAAAGCAGAGCGCCCTCAGTCCACCC 1146
Qy      341 CysAlaAspSerAlaThrSerGluAlaAlaGlyIleuArgProValHisProAlaSerVal 360
Db      1147 TGTGCCACTCTCTACTCAGAGGAGCAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1206
Qy      361 ProSerValProSerValGlnProSerLeuGluAspSerProLeuValGlnAlaLeu 380
Db      1207 CCAGCGTCCCGCAGCGCGCAGCGCTGCTTTAGAGACAGCCCGCTGTACAGCGCTC 1266
Qy      381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuValArgSerPhe 400
Db      1267 AGTGGCTGCACCTGTCCGACAGAGGCTGGAAGACAGCAATAAAGCTCAAGCGTTCCTTC 1326
Qy      401 SerLeuAspIleuLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
Db      1327 TCTCTGATATTCAAATCATGTTCAATATTCAGCAGCAATGACATCTTCAATGCGCTTC 1386
Qy      421 SerSerSerGluAspAlaLeuGluTyrTyrTyrPheProSerThrThrLeuAspGlyThrAsn 440
Db      1387 TCCTCATCAAGAAATGCTTTGGAACTACTACAAACTTCCACTCTCGATGAGGACCAAC 1446
Qy      441 LysLeuGlnPheSerProValGlnGluLeuSerGluIleuThrProGluThrSerPro 460
Db      1447 AAGCTATGCGCAGTTCTCCCTGTTCAAGAACTATCGGAGAGACTCCGCAACCAAGTCT 1506
Qy      461 AspLysGluGluAlaSerIleProLysLysLeuGluThrAlaArgProSerAspSerGln 480
Db      1507 GATAGAGAGAACCGACATCCCAAGAACCTCGAGCCGCGCTTACAGACAGCAG 1566
Qy      481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db      1567 AGCAGCGATTGCAATTCGTCAGAACGACGAGATGGCACCGCCCAAGAGTCCCTTTTA 1626
Qy      501 SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
Db      1627 TCTCCACTGATCGAAGTGGAGGTGAGAGACATTAACACACAGCTTCCTTTTGGC 1686
Qy      521 LeuSerThrSerGlnGlnHisLysLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrHis 540
Db      1687 CTTTCCACCGACGACGACCTCACCAATCTGCTGGCCCTTAAGGGCTGGCAC 1746
Qy      541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrIlePheAla 560
Db      1747 TCGGATATCTGGGCCCCCGACACTCTACCCCTTCCCTGACCGAGACTGATTTTGGC 1806
Qy      561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyIleSerAlaSerTyrSer 580
Db      1807 ACAGAGTCTCACAACCTTACTGCTGCTCAGCCATTAACGAGGAGAGTGCAGTACTCT 1866
Qy      581 AlaTyrSerGlySerGlnLeuProThrCysGlyAspGlnAlaTyrSerValAlaArgArg 600
Db      1867 GCGTACGCTGACGCGCGCTGCCACTTGGAGAACCAATCTATCTGTGCGCAGCGCG 1926
Qy      601 GlnLysProSerAspArgAlaAspSerArgSerTrpHisGluGluSerProPheGlu 620

```

```

Db      1927 CAGACGCAAGTACAGAGACTGACGCGCGAGCTGGCATGAGAGACCCCTTTGAA 1986
Qy      621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
Db      1987 AAGCAGTTTAAAGCAGAGCTGCAAAATGGAATTTGGAGAGCATCATGTCAGAGAAC 2046
Qy      641 ArgSerArgGluGluLeuGluLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db      2047 AGGTACGGGAGAGAGCTGGGGAAGTGGCGATGATGCTTTCGGGACGATGAA 2106
Qy      661 IleIleGluValSer 665
Db      2107 ATCATTGAGGTCTCC 2121

RESULT 5
AEN59704
ID AEN59704 standard; cDNA; 3104 BP.
XX
AC AEN59704;
XX
DT 28-JUN-2002 (first entry)
XX
DE Novel human coding sequence SEQ ID NO: 115.
XX
KW Human; antihaemic; vulnerary; antiinflammatory; immunomodulator;
KW antifertility; cerebroprotective; cyrostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX
OS Homo sapiens.
XX
PN MO200222660-A2.
XX
PD 21-MAR-2002.
XX
PF 10-SEP-2001; 2001MO-US026015.
XX
PR 11-SEP-2000; 2000US-00659671.

(PHYSE-) HYSEQ INC.
XX
PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao Qa, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Dermanac RT;
XX
DR WPI, 2002-292408/33.
XX
PT P-PSDB; ABB97291.
XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.
XX
PS Claim 1; SEQ ID NO 115; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haemtopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention
XX
SQ Sequence 3104 BP; 845 A; 772 C; 718 G; 769 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,04e-234 Length: 3104
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: Gaps: 0

```

US-10-029-345A-109 (1-665) x ABRN59704 (1-3104)

QY 1 MetalAlaHisGluMetCIIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 Db 159 ATGGCCCATGAGATGATTTGAACTCAAAATGTTATACAGAGGTTGGTGGCTCTGGTGGAA 218
 QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 Db 219 AGTGAACCGAAGAAAGTGTCTAATTGATAGCCGCCCATTTGTGGAAATACATACATCC 278
 QY 41 HisIleLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
 Db 279 CACATTTTGGAAACCATTAATATCACTCTCAAGCTTATGACGAAAGGTTGCAACAG 338
 QY 61 AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
 Db 339 GACAAAGTGTATTATACAGAGCTCATCCAGCATTCAGCGAAACATAGGTGACATTGAT 398
 QY 81 CysSerGlnIleValValValValValValValValValValValValValValValVal 100
 Db 399 TGCAGTCAGAGGTTGTATTTATACATCAAGCTCCCAAGATGTTGCTCTCTCTCTCA 458
 QY 101 AspCysPheLeuThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
 Db 459 GACTGTTTCTCAGCTGATCTTCTGGGTAACTGGAGAAAGCTTCACTGTTCACTCG 518
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
 Db 519 CTTCAGAGGTTGGTTTCTGAGATTTCTCTGATGTTTCTCTGCTCTCTGCTGAGAAATCC 578
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 Db 579 ACTCAGTCCCTTACCTGATTCATTCAGACTTTCAGACTTTCAGACTTTCAGACTTTC 638
 QY 161 ArgIleLeuProAsnLeuIleGlyLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
 Db 639 CGAATTTCTCCCATCTTTATCTTGGCTGCAAGCTGCTCTCAACAGAGGCTGATG 698
 QY 181 GlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerIleThrCysProLysProAspPhe 200
 Db 699 CAGCAGATGGAGATGGTATATGTTAAATGCCAGCAATACCTGCTCAAGGCTGACTT 758
 QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
 Db 759 ATCCCGAGTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 818
 QY 221 ProIleLeuAspLysSerValAspPheIleGluLysAlaLysAsnGlyCysVal 240
 Db 819 CCGTGGTGGACAAATCAGTATGATTGAGAAAGCAAGCTTGTGAGAAATTTTG 878
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAla 260
 Db 879 CTATGTCACATGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 938
 QY 261 LysArgMetAspMetSerLeuAspGlnAlaIleArgPheValLysGluLysAspProThr 280
 Db 939 AAGAGAGTGAACATGCTTTTATGATGAAGCTTACAGATTGTGAAGAAAGAAAGCTTACT 998
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIleGluLysIleLysAsn 300
 Db 999 ATATCTCCCACTTCANTTTTCTGGCCCAACTCTGAGCATGAGAAAGAAAGTAAAGAAC 1058
 QY 301 GlnThrGlyLysAspGlyProLysSerLysLeuLysLeuLeuLeuLeuLeuLeuLeu 320
 Db 1059 CAGACTGGAGATCAGGGCCCAAGAGCAACTCAAGCTGCTGCACTGGAGAAAGCAAT 1118
 QY 321 GluProValProAlaValSerGluGlyGlyLysSerGlnThrProLeuSerProPro 340
 Db 1119 GAACCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1178
 QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyIleArgProValHisProAlaSerVal 360

Db 1179 TGTCCGACTCTGCTACCTCAGAGGCAAGAAAGCCCGCTGCATCCCGCAGGCTG 1238
 QY 361 ProSerValProSerValGlnProSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1380
 Db 1239 CCAGAGCTCCAGAGCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1298
 QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
 Db 1299 AGTGGGCTGCACCTGTCTCCGAGACAGGCTGGAAGACGCAATAGCTCAAGGCTTCTTC 1358
 QY 401 SerLeuAspIleLysSerValSerIleSerAlaSerMetAlaIleSerLeuIleGlyPhe 420
 Db 1359 TCTCTGATATCAATCAATCAGTTTATATTCAGCCAGATGCAAGCTTCTTACATGCTTC 1418
 QY 421 SerSerSerGluAspAlaLeuGluIleValIleValIleValIleValIleValIleVal 440
 Db 1419 TCCTCATCAGAAAGTGTGTTGAAATCTCAAACTTCCACTACTGGAATGGACCAAC 1478
 QY 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
 Db 1479 AAGCTATGCCAGTTCTCCCTGTTGAGAACTATCGAAGCACTCCGAAACAGATCCT 1538
 QY 461 AspLysGluGlnLysSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 Db 1539 GATTAAGAGAAAGCCAGCATCCCAAGAGCTGACAGCCGACAGGCTTCAAGACGCAAG 1598
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 Db 1599 AGCAACGATTCATTCGCTGAGAACAGACAGACAGCTGCAAGGCTCCCTTTTAA 1658
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnIleThrIleSerPheLeuPheGly 520
 Db 1659 TCTCCACTGCATCGAAGTGGAGGCTGAGGACAAATTAACAACAGCTTCTTTTCGGG 1718
 QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyIlePhe 540
 Db 1719 CTTCACACAGGCAAGCAGACCACTCAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTG 1778
 QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerIlePheAla 560
 Db 1779 TCGGATATCTTGGCCCCCCCCAGACCTTACCTCCCTGCTGCTGCTGCTGCTGCTG 1838
 QY 561 ThrGluSerSerHisPheIleSerAlaSerAlaIleIleGlyIleSerAlaSerIleSer 580
 Db 1839 ACAGAGTCTCACTCTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1898
 QY 581 AlaIleSerCysSerGlnLeuProThrCysGlyAspGlnValIleSerValArgArgArg 600
 Db 1899 GCTTACAGCTGCAAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1958
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerIleGluGluSerProPheGlu 620
 Db 1959 CAGAGAGCAAGTACAGACTGACTCGGGGAGACTGCAATGAGAGAGGCCCCCTTGAA 2018
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
 Db 2019 AAGCAATTTAAACGCAAGAGCTGCAAAATGGAATTTGAGAGACATCATGTGAGAAC 2078
 QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerArgLysMetGlu 660
 Db 2079 AGGTCAACGGGAAGACTGGGAAAGTGGCAGTCAAGTCAAGCTTTTGGGCAAGCATGAA 2138
 QY 661 IleIleGluValSer 665
 Db 2139 ATCATTTAGAGGTCTCC 2153

RESULT 6
 ABR47596
 ID ABR47596 standard; CDNA; 3496 BP.
 XX ABR47596;
 AC
 DT 02-JUL-2002 (first entry)

XX US2002034807-A1.
 PN
 XX
 PD 21-MAR-2002.
 XX
 XX 23-MAR-2001; 2001US-00816494.
 PF
 XX
 PR 24-MAR-2000; 2000US-0191858P.
 XX
 XX (MEYE/) MEYERS R A.
 PA
 XX
 PI Meyers RA;
 XX
 XX WPI: 2002-351088/38.
 DR
 DR P-PSDB; AAU79929.
 XX
 XX
 PT New nucleic acids, designated 38692 and 21117, encoding dual specificity
 PT phosphatases for treating cell proliferation and differentiation
 PT disorders including hematopoietic and erythroid-related disorders and
 PT cancers.
 XX
 XX
 PS Claim 2; Fig 1; 76pp; English.
 XX
 CC The present invention relates to new nucleic acids designated 38692 and
 CC 21117 encoding dual specificity phosphatase family members. The nucleic
 CC acid, polypeptide encoded by it, and antibody specific for the
 CC polypeptide may be used to diagnose and treat haematopoietic-related
 CC disorders such as leukemias and autoimmune diseases, erythroid-related
 CC disorders such as anaemias and erythrocytosis, liver-related disorders,
 CC and cancers, particularly of the breast, colon, adipose, prostate and
 CC lung. The present nucleic acid sequence encodes the human dual
 CC specificity phosphatase 21117 protein of the invention, as described
 CC above
 XX
 SQ Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.:
 Score: 1.24e-234 Length: 3544
 Percent Similarity: 3406.00 Matches: 663
 Best Local Similarity: 99.85% Conservative: 1
 Query Match: 99.70% Mismatches: 1
 DB: 6 Indels: 0
 Gaps: 0
 US-10-029-345a-109 (1-665) x ABK49402 (1-3544)
 QY 1 MetAlHisGluMetIleGlyThrGluIleValThrGluArgLeuValAlaLeuLeuGlu 20
 DB 589 ATGGCCCATGAGATGATGGAAGTCAAAATGTTACTGAGAGGTTGGCTCTGCTGGAA 648
 QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 DB 649 AGTGGACCGGAAAAGCTGCTGTAATTATGACCGGCATTGGGAAATCAATACATCTCC 708
 QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerLeuMetLeuValAspArgLeuGln 60
 DB 709 CACATTTTGGAAAGCCATTATATCACTGCTCCAAAGTTATGAGAGGAGGTTGCAACG 768
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
 DB 769 GACAAAGTTAATTACAGAGCTATCCAGCATTCACCGAAACATTAAGGTTGACATTGAT 828
 QY 81 CysSerGlnLysValAlaValTyrAspGlnSerSerGlnAspValAlaSerLeuSer 100
 DB 829 TGCAGTCAGAGGTTGATTGATTCGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 888
 QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
 DB 889 GACGTTTTTCTACTGTACTTCTGGGTAACTGAGAAAGACCTCAACTCTGTTCACTG 948
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
 DB 949 CTTCAGAGGTGGTTCTGCTGATTCCTGTTTCCCTGGCCTCTGAGAGGAAATCC 1008

QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 1009 ACTTAGTCCCTACTGATCTTCTCAGCTTGTCTTACTCTGTGGCAATGGGCGCAACC 1068
 QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuLeuLysGluLeu 180
 DB 1069 GAAATCTTCCCAATCTTATCTTGGCTGCAGAGATGCTCTCAACAGAGCTGATG 1128
 QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
 DB 1129 CAGCAGATGGGATTTGGTTAAGTTTAAAGCCAGCAATACCTGTCCAAAGCCTGACTT 1188
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
 DB 1189 ATCCCGAGTCTCAATTTCTGCGTGCCTGTGATGACAGCTTTGTGTAGAAAATTTTG 1248
 QY 221 ProTyrLeuAspLysSerValAspPheIleGluValAlaLysAlaSerAsnGlyCysVal 240
 DB 1249 CCGTGTGGACAAATCAGTACATTTCAATTGAGAAAAGCAAAAGCCTCCAAATGATGTT 1308
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 DB 1309 CTAGTGACATGTTTACTGAGATCTCCGCTCCGCCACATCGCTATGCTTACATCATG 1368
 QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
 DB 1369 AAGAGATGACATGCTTTATATACATTAACCTTACATTTGGAAGAAAAGAACTTACT 1428
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
 DB 1429 ATATCTCCAACTTCAATTTCTGGCCCACTCTGACATGAGAGAAAGATTAGAAC 1488
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGluLysProAsn 320
 DB 1489 CAGACTGGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGACCTGGAGAAAGCCAAAT 1548
 QY 321 GluProValProAlaValAsnSerGluGlyGlnLysSerGlnThrProLeuSerPro 340
 DB 1549 GAACCTGCTCCCTGCTGCTCTGAGAGGTGACAGAAAAGCAGCCCTCAGTCCACCC 1608
 QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGluArgProValHisProAlaSerVal 360
 DB 1609 TGTGCGGACTGTGCTACTCAGAGGACAGAGCAAAAGCCGTGATCTCCGACGCTG 1668
 QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
 DB 1669 CCCAGCTGCCCAAGCGGACGCTGCTGTGTTAAGAGACAGCCCTGTTACAGCGCTC 1728
 QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
 DB 1729 AGTGGGCTGCACCTGTCCGACAGCGCTCGAAGACAGCAATTAAGCTTAAGCTTCTTC 1788
 QY 401 SerLeuAspIleLysSerValSerTyrSerAlaSerMetValAlaSerLeuHisGlyPhe 420
 DB 1789 TCTCTGATATCAAACTATTCATATTCAGCCAGCATGAGAGATCTTACATGGCTTC 1848
 QY 421 SerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
 DB 1849 TCCTCATCAGAAAGATGCTTGGAAATACCAAACTTCCACTACCTGATGGAGCAAC 1908
 QY 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGlnThrProGluThrSerPro 460
 DB 1909 AAGCTATGCAATTTCTCCCTGTTCAGAACTATTCAGAGAGACTCCCGAACAAGTCT 1968
 QY 461 AspLysGluGluAlaSerIleProLysLeuLeuGlnThrAlaArgProSerArgAspSerGln 480
 DB 1969 GATTAAGAGAAAGCAGCATCCCAAGACTCGAAGCCCGAGGCTTTCAGACACCG 2028
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 DB 2029 AGCAAGCATTCATTCGTTGGTGAACCAAGACAGAGTGCACCGCCAGAGGTCCCTTTA 2088

QY 501 SerProteinHsArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
 DB 2089 TCTCCACTGATCGAAGTGGAGCGTGGAGCAATTAACACACAGCTTCTTTGGC 2148
 QY 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuGlyTyrPheHis 540
 DB 2149 CTTTCCACGACGACGACATCAAGAGTGTGCTGCGCTGGGCTTAAGGGCTGGCAC 2208
 QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
 DB 2209 TCGGAATATCTGGCCCCCGACCTTAACCCCTTCCCTGACACGACGCTGTATTTTGGC 2268
 QY 561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
 DB 2269 ACAGAGTCTCAACCTTCTACTGCTCTGCTCAGCCATCTACGAGGAGGAGTCACTTACTCT 2328
 QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValAlaArgArg 600
 DB 2329 GCCTACAGCTGACGACGACGCTGCCACTTGGCGAGACCAAGTCTATCTGTGGCAGGCGG 2388
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrPheGluGluSerProPheGlu 620
 DB 2389 CAAAGGCAAGTACACAGAGCTGACTCGCGCGGAGCTGATGAGAGAGACCCCTTTGAA 2448
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
 DB 2449 AACGAGTTTAAACGACAGAGCTGCCAATGGAATTGGAGAGCATCATGTCCAGAGAAC 2508
 QY 641 ArgSerArgGlnGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 DB 2509 AGGTCAAGGAGAGGCTGGGGAAGTGGGAGTCAAGTCACTTTTGGGCGACAGATGAA 2568
 QY 661 IleIleGluValSer 665
 DB 2569 ATCATTTAGAGTCTCC 2583
 RESULT 9
 ID ABK14474 standard; cDNA; 3766 BP.
 AC ABK14474;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Human protein phosphatase 7 (PP7) cDNA sequence.
 XX
 KW Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
 KW acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
 KW Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
 KW dementia; Parkinson's disease; developmental disorder; Down's syndrome;
 KW cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
 KW melanoma; myeloma sarcoma; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 536..2535
 FT /tag= a
 FT /product= "Protein_phosphatase_7 (PP7)"
 XX
 PN MO200210363-A2.
 PD 07-FEB-2002.
 XX
 PF 26-JUL-2001; 2001MO-US023716.
 XX
 PR 28-JUL-2000; 2000US-0221679P.
 PR 03-AUG-2000; 2000US-0223272P.
 PR 10-AUG-2000; 2000US-0224309P.
 PR 18-AUG-2000; 2000US-0226728P.
 PR 30-AUG-2000; 2000US-0229254P.
 PR 08-SEP-2000; 2000US-0231366P.
 XX

PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Teng YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;
 PI Stewart EA, Gandhi AR, Patterson C, Lee EA, Hafalia AJL, Lu DAM,
 PI Tribouley CW, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;
 PI Walla NK, Kearney L;
 XX
 DR WPI; 2002-188735/24.
 DR P-PSDB; AAU75789.
 XX
 PT New protein phosphatases, useful for diagnosing, treating or preventing
 PT immune system disorders (e.g. Crohn's disease), neurological disorders
 PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
 PT cancers).
 XX
 PS Claim 5; Page 114-115; 117pp; English.
 XX
 CC The present invention relates to a new polypeptide, a naturally occurring
 CC amino acid sequence at least 95 % identical to it, a biologically active
 CC fragment of it or an immunogenic fragment of it. The polypeptides,
 CC polynucleotides, agonists and antagonists are useful for diagnosing,
 CC treating or preventing disorders associated with aberrant expression of
 CC protein phosphatases (PP), particularly immune system disorders e.g.
 CC acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
 CC asthma or Crohn's disease, neurological disorders e.g. epilepsy,
 CC Huntington's disease, dementia or Parkinson's disease, developmental
 CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
 CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
 CC or sarcoma. The present nucleic acid sequence encodes human protein
 CC phosphatase 7 (PP7) which is one of several human protein phosphatases
 CC (AAU75783-AAU75792) of the invention
 XX
 SQ Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,346-234 Length: 3766
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-345A-109 (1-665) x ABK14474 (1-3766)
 QY 1 MetalHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
 DB 538 ATGGCCCATGAGATGATTTGGAACTCAAAATGTTACTGAGAGGTTGGCTGTGCGGAA 597
 QY 21 SerGluThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 DB 598 AGTGGAAACGAAAGATGCTGCTTAATGATGATGAGCGCATTTTGGAAATACATACATCC 657
 QY 41 HisIleLeuGluValIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
 DB 658 CACATTTTGGAGCCATTAAATACATCTCCAACTTATGAGGAGGAGGTTGCAACAG 717
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAlaPheAsp 80
 DB 718 GACAAAGTGTATATTACAGAGCTCATTCAGCATTCAGCGAAACATAGAGTTGACATTGAT 777
 QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB 778 TGCAGTCAGAGAGGTATAGTTTACATCAAGATCCCAAGATTTTCTCTCTCTTCA 837
 QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
 DB 838 GACTGTTTCTCACTGATCTTCTGGGTAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 897
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlyGlyLysSer 140
 DB 898 CTTCAGAGTGGGTTGCTGATGTTCTCTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160

```

Db 958 ACTTAGTCCCTACTGATTTCTCAGCCTTCTACTCTGTCCAACTTGGCCACAC 1017
Qy 161 Arg1LeuProAlaLeuTyrLeuGlyCysGlnArgSpAlaLeuAlaLeuGlyLeu 180
Db 1018 CGAATTTCTCCAAATCTTATCTTGTGCTGCAGAGATGTCTCTCAACAAGAGCTGAG 1077
Qy 181 GlnGlnAmGlyIleGlyTyrValLeuAlaSerTyrThrCysProLysProAlaPhe 200
Db 1078 CAGCAGATGGATTTGGTTATGTGTAAATCCAGCAATACCTGTCCAAAGCCTCAGCTT 1137
Qy 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlyLeuValLeu 220
Db 1138 ATCCCCAGATCTCAATTTCTGCTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 1197
Qy 221 ProTyrLeuAspLysSerValAspPheIleGlyLysValAlaLysAlaSerAmGlyCysVal 240
Db 1198 CCGTGTGTGACAAATCAGTAAATTTTCAATTGAAAGCAAAAGCCTCCAAATGAGTGTGTT 1257
Qy 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db 1258 CTAGTCACTGTGTAGTGTGAGATCTCCGCTCCGCCCATCGCTATCGCTTACATCATG 1317
Qy 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlyLysArgProThr 280
Db 1318 AAGAGGATGACATGCTTATGATGACCTTACAGATTTGTGAAGAAAGAAAGACCTACT 1377
Qy 281 IleSerProAlaPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLysLysIleLysAsn 300
Db 1378 AATATCTCAAACTTCAATTTTCTGAGCCCACTCTGACATATGAAAGAAAGTATTAAGAC 1437
Qy 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysIleLeuGlyLysProAla 320
Db 1438 CAGACTGAGCATCAGGGCCAAAGACAACTCAAGCTGTGACCTGAGAGAGCAAT 1497
Qy 321 GluProValProAlaValSerGlyGlyGlnLysSerGlnThrProLysSerProPro 340
Db 1498 GAACCTGCTCCCTGTCTCTCAGAGGGTGCAGAGAAAGCAAGCCCTTCAGTCAACC 1557
Qy 341 CysAlaAspSerAlaThrSerGlnAlaGlyGlnArgProValHisProAlaSerVal 360
Db 1558 TGTGCCGACTCTGCTCAGAGGACAGAGCAAAAGCCCTGCTATCCGCCAGCTC 1617
Qy 361 ProSerValProSerValGlnProSerLeuGlnAspSerProLeuValGlnAlaLeu 380
Db 1618 CCCAGGCTGCCAGCGTGCAGCGCTGTAGAGGACAGCCGCTGTACAGCCCTC 1677
Qy 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAlaLeuLysArgSerPhe 400
Db 1678 AGTGGGCTGACCTGTCCGACAGAGCTGCAGAGCAAGCAATTAAGCTTCAGCCCTTCC 1737
Qy 401 SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
Db 1738 TCTCTGGATATCAAAATCAGTTTCAATTCACGCAAGTGGCAGATCTTACATGCTTC 1797
Qy 421 SerSerSerGlnAspAlaLeuGlyTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
Db 1798 TCCCTCATCAGAAAGATGCTTGAATATCTCAAACTTCCACTCTCTGATGGAGCAAC 1857
Qy 441 LysLeuCysGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
Db 1858 AAGCATGCCAGTTCTCCCTGTTCAGAGAACTATCGAGAGCAGCTCCGAAACAGACTCT 1917
Qy 461 AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspGln 480
Db 1918 GATAGAGAGAGAGCAGCATCCCAAGAACTCAAGCCGCTTCAGACAGCCAG 1977
Qy 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db 1978 AGCAGAGCATTTGATGTGTGACAGAACAGAGTGTGACCGCCAGAGATCCCTTTTA 2037
Qy 501 SerProLeuHisArgSerGlySerValGlnAspAlaThrThrSerPheLeuPheGly 520

```

```

Db 2038 TCTCCAGCTGCATCAGATGGAGCGGTGAGAGCAATTACACACAGCTTCTTTTGGC 2097
Qy 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysGlyTyrHis 540
Db 2098 CTTTCCAGCAGCAGCAGACACCTTCAAGAAATCTGTGCTGCGCTGAGGCTTAAAGGCTGGC 2157
Qy 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
Db 2158 TCGAATATCTTGGCCCCCAGACCTTACCTTCCCTTCCGACAGACCTGATATTTTGGC 2217
Qy 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSer 580
Db 2218 ACAGAGTCTCACTTCTACTCTGTCTGAGCAATCTTACGAGAGCAGTGTCCAGTTACTCT 2277
Qy 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
Db 2278 GCTTACAGCTGACCCAGCCTGCCCATTTGGAGACCAAGTCTATTTCTGTGCGACAGCG 2337
Qy 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrHisGlnGlnLysSerProPheGlu 620
Db 2338 CAGAAAGCAAGTACAGAGCTGACTGCGAGCTGAGCATGAGAGAGAGCCCTTTGAA 2397
Qy 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
Db 2398 AAGCATTTTAAAGCAGAGAGCTGCCAAATGGAATTTGGAGAGAGCATATGTCAGAGAC 2457
Qy 641 ArgSerArgGlnGlyLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db 2458 AAGCTCAGGAGAGAGCTGGGAAAGTGGAGCTGACGTCTTGTGCGGACGATGAA 2517
Qy 661 IleIleGlnValSer 665
Db 2518 ATCATTTAGAGTCTCC 2532

```

```

RESULT 10
ABN83966
ID ABN83966 standard; DNA; 4790 BP.
XX
AC ABN83966;
XX
DT 06-SEP-2002 (first entry)
XX
DE Human gene sequence #13.
XX
KW Human; brain; tonsil; hippocampus; foetal brain; diagnosis; gene; ds.
XX
OS Homo sapiens.
XX
FH
FT Key Location/Qualifiers
FT CDS 184..2181
FT CDS /*lag= a
XX
PD WO200252005-A1.
XX
PD 04-JUL-2002.
XX
PE 20-DEC-2001; 2001MO-JP011217.
XX
PR 22-DEC-2000; 2000JP-00389742.
XX
PA (KAZU-) KAZUSA DNA RES INST FOUND.
PA (CELE-) CELESTAR LEXICO-SCI LTD.
PI Ohara O, Nagase T, Nakajima D;
PI XX
DR WPI; 2002-500762/53.
DR P-PSDB; ABB97946.
XX
PT Genes and their expression products cloned from human cDNA libraries for
PT treatment and diagnosis of diseases associated with their expression.
XX
PS Claim 1(a); Page 111-117; 238bp; Japanese.
XX

```

CC The invention relates to DNA encoding polypeptides directly cloned from
CC cDNA libraries originating in adult whole brain, human tonsil, human
CC adult hippocampus and human foetal whole brain. Polypeptides and
CC polynucleotides of the invention may be used in the investigation of
CC differential expression of the DNA sequences in normal subjects and
CC disease patients. They may also be used in the production of antibodies,
CC oligonucleotide probes and DNA chips for diagnosis and identification of
CC drugs for treatment of diseases with which the DNA sequences are
CC associated. The sequences given in records ABN83954-ABN83984 represent
CC human gene sequences of the invention

XX Sequence 4730 BP; 1293 A; 1064 C; 1152 G; 1281 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.85e-234	Length:	4730
Score:	3406.00	Matches:	663
Percent Similarity:	99.85%	Conservative:	1
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	99.65%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345a-109 (1-665) x ABN83966 (1-4790)

Qy 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
Db 184 ATGGCCCATGAGATGATGGAACTCAAAATTGTTACTGAGAGTTGGCTGCTGCGAA 243
Qy 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyraanthSer 40
Db 244 AGTGAACCGAAAGAGTGTGCTAATGATGACCGGCACTTGGGAATCAATACATCC 303
Qy 41 HisIleLeuGlnIleAlaLeuIleAsnGlySerIleuMetIleuAspArgLeuGln 60
Db 304 CACATTTGGAGCCATTAAATATCACTGCTCCAAAGCTTATGAAGCAAGTGGCAAG 363
Qy 61 AspLeuValLeuIleThrGluLeuIleGlnHisSerAlaIleValIleAsp 80
Db 364 GACAAAGTGTAAATACAGAGCTATCCACATTCACGAAACAAAGTGGACATTGAT 423
Qy 81 CysSerGlnIleValValIleValIleAspGlnSerSerGlnIleValIleSer 100
Db 424 TGCAGTCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 483
Qy 101 AspCysPheLeuThrValIleLeuGlyValLeuGlnIleuSerPheAsnSerValIle 120
Db 484 GACTGTTTCTCACTGATCTTGGGTAACCTGGAAGAGCTTCACTGTTCACTG 543
Qy 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuGlyIleuSer 140
Db 544 CTTCGAGTGGGTTTGGTGTGATGATCTCTGTTGTTTCCCTGCTGTGAAAGAAATCC 603
Qy 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyPro 160
Db 604 ACTGATGCTCTACCTGATCTTCTGAGCTTGTGCTTACCTGTGCAACATTTGGCC 663
Qy 161 ArgIleLeuProAsnLeuIleValIleGlyCysGlnArgAspValIleuLeuIle 180
Db 664 CGAATCTTCCCAATCTTATCTTGTGCTGACGAGATGCTCTCAACAAGAGCTGATG 723
Qy 181 GlnGluAsnGlyIleGlyIleValIleuAsnAlaSerTyThrCysProIleProAsp 200
Db 724 CACGAAATGGATGGATTGTGTGTTAAATGCAAGCAATCTGTTCAAGAGCTGACTTT 783
Qy 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluValIle 220
Db 784 ATCCCGAGTCTCATTTCTGCGTGTCTGTGATGACAGCTTTTGTGAATAATTTTG 843
Qy 221 ProTrpLeuAspIleSerValAspPheIleGluValAlaValAsnArgGlyCysVal 240
Db 844 CCGGTGTGGAACAATGATGATTTCTATGGAAGCAAAAGCTTCAATGATGATGTT 903
Qy 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlMet 260

Db 904 CTAGTCACTGTTTACGTGGATCTCCCGCTCCGCCACCATCGCTATCGCTATCATCATG 963
Qy 261 LysArgMetAspMetSerLeuAspGluAlaTyraArgPheAlaIleGlyIleuAspProThr 280
Db 964 AAGAGATGAGACATGCTTTAGATGAGAGCTTACAGATTTGTGAAAGAAAGAACTACT 1023
Qy 281 IleSerProAsnPheAsnPheLeuGlyIleuLeuAspTyrlGluValIleValAsn 300
Db 1024 ATATCTCCAAACTTCAATTTTGTGGCCCAATCTTGACATATGAAAGAAATTAAGAAC 1083
Qy 301 GlnThrGluAlaSerGlyProIleuSerIleuLeuIleuLeuIleuIleuIleuIleu 320
Db 1084 CAGATGAGACATCAGAGCCCAAGAGCAAACTCAAGCTGCTGACCTGAGAGAGCCAAAT 1143
Qy 321 GlnProValProAlaValSerGluGlyValGlnIleuSerGluThrProLeuSerProPro 340
Db 1144 GAACCTGTCTCTGCTCTCTCAAGAGGTGACAAAGAGAGAGAGAGAGAGAGAGAGAG 1203
Qy 341 CysAlaAspSerAlaThrSerGluAlaIleGlyIleuArgProValHisProAlaSerVal 360
Db 1204 TGTGCCGACTCTGCTACCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1263
Qy 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
Db 1264 CCAGAGGTGCGCAGCGTGCAGCGCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1323
Qy 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIleuValAspSerPhe 400
Db 1324 AGTGGCTGCACTGCTCCGAGACAGAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAG 1383
Qy 401 SerLeuAspIleLeuSerValSerTyrlSerAlaSerMetAlaIleSerLeuHisGlyPhe 420
Db 1384 TCTTGATATCAAAATCAAGTTTATATTCAGACGATGACAGATCTTACATGCTTC 1443
Qy 421 SerSerSerGlnAspAlaLeuGluTyrlTyrlPProSerThrThrLeuAspGlyThrAsn 440
Db 1444 TCTCATCAGAAAGATGCTTGGAAATGATCAAACTTCCATCTGATGAGAGAGAGAG 1503
Qy 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
Db 1504 AAGCTATGCAAGTTCCTCCCTGTTAGAGAACTTTCAGAGAGAGAGAGAGAGAGAG 1563
Qy 461 AspLeuGluGlnIleSerIleProIleuValLeuGlnIleThrAlaArgProSerAspSerGln 480
Db 1564 GATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1623
Qy 481 SerIleArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
Db 1624 AGCAAGCATGATTCGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1683
Qy 501 SerProLeuHisArgSerGlySerValGluAspAsnTyrlHisThrSerPheLeuPheGly 520
Db 1684 TCTCCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1743
Qy 521 LeuSerThrSerGlnGlnIleuThrIleuValIleuGlyIleuValIleuValIleuVal 540
Db 1744 CTTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1803
Qy 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrlTyrlPheAla 560
Db 1804 TGGATATCTGGCCCCAGAGCTTACCCCTTCCCTGACAGAGAGAGAGAGAGAGAGAG 1863
Qy 561 ThrGluSerSerHisPheTyrlSerAlaIleTyrlGlyIleSerAlaSerTyrlSer 580
Db 1864 ACAGAGTCTCTCACTTCACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1923
Qy 581 AlaTyrlSerCysSerGlnIleuProThrCysGlyAspGlnValTyrlSerValArgAspArg 600
Db 1924 GCTTACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1983
Qy 601 GlnIleProSerAspArgAlaAspSerArgArgSerTyrlHisGluIleuSerProPheGlu 620
Db 1984 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2043

QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyCysUterSerLeuSerGluasn 640
 DB 2044 AAGCAGTTTAAACCCAGAGCTGCCAATATGAAATTTGGAGAGCATCATGTCCAGGAAC 2103
 QY 641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 DB 2104 AGGTCACGGGAGAGGCTGGGGAAAGTGGCAGTCACCTTACGCTTTTCGGGACAGATGAA 2163
 QY 661 IleIleGluValSer 665
 DB 2164 ATCATTGAGGTCTCC 2178
 RESULT 11
 ABV20833
 ID ABV20833 standard, cDNA; 5145 BP.
 AC ABV20833;
 XX 13-SEP-2002 (first entry)
 DT 13-SEP-2002 (first entry)
 DE Human prostate expression marker cDNA 20824.
 XX
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MIL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 3419; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring whether
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (g)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
 Alignment Scores:
 Pred. No.: 2.03e-234 Length: 5145
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1

Query Match: 99.65% Indels: 0
 DB: 5 Gaps: 0
 US-10-029-345a-109 (1-665) x ABV20833 (1-5145)
 QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluValAlaLeuGlu 20
 DB 589 ATGGCCCATGAGATGATGGAATCTCAATTTGTTACTGAGAGGTGGTGGCTGCTGGAA 648
 QY 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluThrAsnThrSer 40
 DB 649 AGTGGACGCAAAAGAGCTGCTTAATTGATAGCCGGCATTTTGGAAATACATACATCC 708
 QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerLysLeuMetLysArgLeuGln 60
 DB 709 CACATTTTGGAAAGCCATTAATATCACTGCTCCAAACCTTATGAGAGGAGTTGCAACG 768
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysValAspIleAsp 80
 DB 769 GACAAAGTTTAATTACAGAGCTATCCAGCATTCAGCAAAACATPAGGTGACATTGAT 828
 QY 81 CysSerGlnLysValValIleThrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB 829 TGCAGTCAGAGGTTGATGTTACATCAAAAGCTCCAAAGATGTTGCTCTCTCTCA 888
 QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeu 120
 DB 889 GACTGTTTCTCACTGTACTTCTGGGTAACTCGAGAAAGCTTCACTGTTCACTG 948
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
 DB 949 CTTCAGAGTGGGTTTGTGAGTTCCTGCTGTTGTTCCCTGGCCCTCTGGAAGAAATCC 1008
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 1009 ACTCTAGTCCCTACCTGATTTCTCAGCTTGTCTTACTGTTGCAACATTTGGCCCAACC 1068
 QY 161 ArgGlyLeuProAsnLeuThrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuIle 180
 DB 1069 CGAATTTCTCCAAATCTTATCTGGCTGCAGAGATCTCTCAACAGAGGCTGAGT 1128
 QY 181 GlnGlnAsnGlyIleGlyThrValLeuAsnAlaSerThrCysProLysProAspPhe 200
 DB 1129 CAGCAGATGGAGTGGTATGTTAATGCCAAGAAATCTGTCAAAAGCTGACATT 1188
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
 DB 1189 ATCCCGAGTCTCATTTCTGCGTGGCTGTGAATGACAGCTTTGTGAGAAATTTTG 1248
 QY 221 ProThrLeuAspLysSerValAspPheIleGluLysAlaLysSerAsnLysCysVal 240
 DB 1249 CCGTGGTTGGACAAATCAGTAAATTTCATTGAGAAAGCAAAAGCTCCAAAGAGATGTT 1308
 QY 241 LeuValHisCysLeuValIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 DB 1309 CTAGTGCATGTTTACTGGAATCTCCGCTCCGACCAATCGTATCGCTACATCATG 1368
 QY 261 LysArgMetAspLysSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
 DB 1369 AAGAGGATGAGATGCTTATGATTAACCTTACAGATTGTAAGAAAGAAAGAAAGCTTACT 1428
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLysIleLysAsn 300
 DB 1429 ATATCTCCAAACTTCAATTTCTGCGCCACCTCTCGATATGAGAGAGATTAAGAAC 1488
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlyLysProAsn 320
 DB 1489 CAGACTGAGATCATCGGCGCAAGGCAAACTCAAGCTGCTGCACTGGAGAGCCAAAT 1548
 QY 321 GluProValProAlaValAsnSerGluGlyGlnLysSerGluThrProLeuSerProPro 340
 DB 1549 GAACCTGTCCTGCTGCTCAGAGGTGAGCAGAAAAAGCAGCGCCCTCAGTCCACCC 1608


```

QY 341 CyeAlaaspSerAlaThrSerGluAlaGlyGlnArgProValHisProAlaSerVal 360
DB 1609 TGtGCCACtCTGCTACTGACGAGCAGCAGCAAGAGCCGCTGATCCCGCAGCGTG 1668
QY 361 ProSerValProSerValGlnProSerLeuGlnuApsSerProLeuValGlnAlaLeu 380
DB 1669 CCCAGCGTGCCAGCGGAGCGCGCTGCTGTAGAGGACAGCCGCGTGTACAGCGGCTC 1728
QY 381 SerGlyLeuHisLeuSerAlaApsArgLeuGlnuApsSerValLeuValArgSerPhe 400
DB 1729 AGTGGCGTGCACCTGCTCCGAGCAGGCTGTGAAGACAGCAATATAGCTCAAGCCCTTC 1788
QY 401 SerLeuApsIleLeuSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
DB 1789 TCCTGTGATTCATAATCAGTTTATTCATTCAGCCAGCATGCGACATCTTACATGGCTTC 1848
QY 421 SerSerSerGlnuApsAlaLeuGlnuTyrTyrLeuProSerThrThrLeuApsGlyThrAsn 440
DB 1849 TCCTCATCAGAAAGATGTTGTAATACTACAAACCTTCACATCTCGATGGGACCAAC 1908
QY 441 LysLeuCyAGlnPheSerProValGlnuLeuSerGlnuGlnuThrProGlnuThrSerPro 460
DB 1909 AACCTATGCGAGTTCCTCCCTGTTCAGGAATATCGAGCAGACTCCCGAAACAGTCCCT 1968
QY 461 ApsLysGlnuAlaSerIleProLysLysLeuGlnuThrAlaArgProSerApsSerGln 480
DB 1969 GATTAAGAGAGAACCCAGCATCCCGAAGACCTGACAGCCCGGCTTTACAGACCCAG 2028
QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
DB 2029 AGCAGAGATGATTCGATCGGTACAGACAGCAGCAGTGGCAGCCCGCAGAGTCCCTTTTA 2088
QY 501 SerProLeuHisArgSerGlySerValGlnuApsAsnTyrHisThrSerPheLeuPheGly 520
DB 2089 TCTCCACTGATCGAAGTGGAGCGTGGAGACATTAACACACAGCTTCCTTTTCGCGC 2148
QY 521 LeuSerThrSerGlnuGlnuHisLeuThrLysSerAlaGlyLeuGlnuLeuLysGlyTPHs 540
DB 2149 CTTTCCACCGCAGCAGCAGCCTCACAAATCTGCTGGGCTTTAAGGCTGGCAC 2208
QY 541 SerApsIleLeuAlaProGlnuThrSerThrProSerLeuThrSerSerTyrTyrPheAla 560
DB 2209 TCGGATATCTGGGCCCCCGACCACTTACCCCTTCCTGACAGCAGCTGTATTTTGGC 2268
QY 561 ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSer 580
DB 2269 ACAGAGCTCCACACATCTCTGCTGCTCAGCCATCTACGAGGAGGAGTCCAGTTTACTCT 2328
QY 581 AlaTyrSerCySerSerGlnuLeuProThrCyAGLysApsGlnuValTyrSerValArgArgArg 600
DB 2329 GCCTACAGCTGACAGCAGCTGCCCATCTTGCAGAGACCAAGTCTATTTCTGTGCCAGGCGG 2388
QY 601 GlnuLysProSerApsArgAlaApsSerArgArgSerTyrHisGlnuLysProPheGlnu 620
DB 2389 CAAAGCCAAAGTACAGAGCTGACTCGCGCGAGCGTGCATGAAGAGAGACCCCTTTGAA 2448
QY 621 LysGlnPheLysArgArgSerCyAGLysMetGlnuPheGlyLysSerIleMetSerGlnuAps 640
DB 2449 AACGAGTTTAAACGCGAAGACTCCCAATGAAATTTGGAAGAGCATCATGATGAGAAAC 2508
QY 641 ArgSerArgGlnuGlnuLeuGlyLysValGlySerGlnuSerSerPheSerGlySerMetGlnu 660
DB 2509 AGGTACAGGGAAGAGCTGGGGAAGATGGGAGTCAAGTCTTTTCGGGCAACATGGA 2568
QY 661 IleIleGlnuValSer 665
DB 2569 ATCATGAGGTCTCC 2583

```

RESULT 12
 ABV21080
 ID ABV21080 standard; cDNA; 5145 BP.
 XX
 AC ABV21080;

```

XX 13-SEP-2002 (first entry)
DT Human prostate expression marker CDNA 21071.
XX
DE Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
PF 20-FEB-2001, 2001MO-US005171.
XX
PR 17-FEB-2000; 2000US-018319P.
PR 16-MAR-2000; 2000US-018962P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI, 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 3481; 11750bp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

```

Alignment Scores:
 Pred. No.: 2.03e-234 Length: 5145
 Score: 3406.00 Matches: 663
 Percent Similarity: 99.85% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: Gaps: 0

```

US-10-029-345A-109 (1-665) x ABV21080 (1-5145)
QY 1 MetAlaHisGlnuMetIleGlyThrGlnuIleValThrGluArgLeuValAlaLeuGlnu 20
DB 589 ATGGCCCATGAGATGATGAGAACTCAAAATTTTACTGAGAGGTTGGTGGCTGCTGCGAA 648
QY 21 SerGlyThrGlnuLysValLeuLeuIleApsSerArgProPheValGlnuTyrAsnThrSer 40
DB 649 AGTGAACCGAATAAAGTCTGCTAATTTGATGACCGGCAATTTGTGAATACATACATCC 708
QY 41 HisIleLeuGlnuAlaIleAsnIleAsnCySerLysLeuMetLysArgArgLeuGlnuGlnu 60
DB 709 CACATTTTGAAGCCCTTATATCACTCTCCAAAGCTTATGAAGCAAGGTTTGAACAG 768
QY 61 ApsLysValLeuIleThrGlnuLeuIleGlnuHisSerAlaIleHisLysValApsIleAps 80

```

```

Db      769 GACAAAGCTTATTTACAGAGCTCATCCAGATTCCAGAAACATAGAGTTGACATTGAT 828
Qy      81 CysSerGlnLysValValValValYrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db      829 TGCAGTCCAGAAAGTTGATGTTTACGATCAAAAGCTCCCAAGATCTTCTCTCTCTCA 888
Qy      101 AspCysPheLeuThrValLleuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
Db      889 GACTGTTTCTCACTGACCTTCTGAGGAAACCTGGAAGAGCTTCACTCTGTTCACTG 948
Qy      121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
Db      949 CTGCAAGTGGGTTTGCTGAGTCTCTCGTTGTTCCCTGAGCCTCTGGAAGGAAATCC 1008
Qy      141 ThrLeuValProThrCysLleSerGlnProCysLeuProValAlaAsnLleGlyProThr 160
Db      1009 ACTGATGCCCTTACCTGCACTTCTCAAGCTTCTTACCTGTTGCAACATTGGGCCAAC 1068
Qy      161 ArgLleLeuProAsnLeuLysLeuGlyCysGlnArgAspValLleuAsnLysGluLleu 180
Db      1069 CGAATTCCTCCCAATCTTATCTTGCTGCCAGGAAATGCTCTCAAGAGAGCTGATG 1128
Qy      181 GlnGlnAsnGlyLleGlyTrpValLleuAsnAlaSerTrpThrCysProLysProAspPhe 200
Db      1129 CAGCAGAAATGGATTTGGTTATGTTAATGCCAGCAATACCTGTCCAAAGCTGACTTT 1188
Qy      201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlyLysLleu 220
Db      1189 ATCCCGAGTCTCATTTCTCGTGTGCTGTGAAATGACAGCTTTGTGAAATTTTG 1248
Qy      221 ProTrpLeuAspLysSerValAspPheLleGlyLysAlaLysAlaSerAsnGlyCysVal 240
Db      1249 CCGTGTGTGCAAAATCAAGATTTTCAATTGAGAAAGCAAAAGCTTCCATGGATGTGT 1308
Qy      241 LeuValHisCysLeuAlaGlyLleSerArgSerAlaThrLleAlaLleAlaTrpLleMet 260
Db      1309 CTAGTGCACGTGTTAGCTGGGATCTCCGCTCCGCCACCAATGCTATCCCTACATCATG 1368
Qy      261 LysArgMetAspMetSerLeuAspGluAlaTrpArgPheValLysGlyLysArgProThr 280
Db      1369 AAGAGATGACATGCTTATGATGAAAGCTTACAGATTTGTGAAGAAAGAAAGACTTACT 1428
Qy      281 IleSerProAsnAspPheAsnLleuGlyLleuLeuAspTrpGluLysLysLleLysAsn 300
Db      1429 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCGACTGATGAGAAAGAAATTAAGAAC 1488
Qy      301 GlnThrGlyAlaSerGlyProLysSerLysLysLysLysLysLysLysLysLysLys 320
Db      1489 CAGACTGACATCAGAGGCCAAAGCAAACTCAAGCTGCTGCACTGGAGAGCCAAAT 1548
Qy      321 GluProValProAlaValSerGluGlyLysGlnLysSerGluThrProLeuSerProPro 340
Db      1549 GAACCTGTCTCTGCTGCTCAGAGGGTGAGACAGAAAGAGAGAGAGAGAGAGAGAGAG 1608
Qy      341 CysAlaAspSerAlaThrSerGluAlaAlaGlyLysGlnArgProValHisProAlaSerVal 360
Db      1609 TGTGCGCACTCTGCTTACCTCAGAGCAGAGCAAGAAAGCCCGATCCCGCAGACCGTG 1668
Qy      361 ProSerValProSerValGlnProSerLeuGluAspSerProLeuValGlnAlaLeu 380
Db      1669 CCCAGCGGCCAGCGAGCGAGCGCTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1728
Qy      381 SerGlyLeuHisLysSerAlaAspArgLeuGluAspSerAsnLysLysLysLysLysLys 400
Db      1729 AGTGGCTGCACTGCTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCTTCTCTTC 1788
Qy      401 SerLeuAspLleLysSerValSerTrpSerAlaSerMetAlaLysLysLysLysLysLys 420
Db      1789 TCTCTGATATCAATCAATTCATATTCAGCAGATGGCAGATCTTCAATGAGCTTC 1848
Qy      421 SerSerSerGluAspAlaLeuGlyLysTrpLysProSerThrThrLeuAspGlyThrAsn 440

```

```

Db      1849 TCCCTACAGAAAGATGCTTTGGAAATCTTCAAAACCTTCACTACTCTGATGGAGCAAC 1908
Qy      441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
Db      1909 AACGTATGCAATTTCTCCCTGTTTCAAGAACTATCGAGAGCACTCCCGAAACCACTCT 1968
Qy      461 AspLysGluGluAlaSerLleProLysLysLysLysLysLysLysLysLysLysLysLys 480
Db      1969 GATTAAGAGAAAGCCAGCATCCCAAGAGCTGAGACCCCGAGGCTTTCAGACAGCCAG 2028
Qy      481 SerLysArgLeuHisLysSerValArgThrSerSerSerGlyThrAlaGlnArgSerLleu 500
Db      2029 AGCAAGCAATTCAGATTCGCTGACAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2088
Qy      501 SerProLeuHisArgSerGlySerValGluAspAsnTrpHisPheSerPheLeuPheGly 520
Db      2089 TCTCCACTGACATCGAAGGGAGGGAGGTGGAGAGACAAATTCACACAGCTTCTTTCCGC 2148
Qy      521 LeuSerThrSerGlnGlnHisLysLeuThrLysSerAlaGlyLeuGlyLysLysGlyTrpHis 540
Db      2149 CTTTTCACAGCCAGCAGACACCTTCAAGAGCTGCTGCTGGCTTAAAGGCTGGCAC 2208
Qy      541 SerAspLleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTrpPheAla 560
Db      2209 TCGGATATCTTGGCCCCCGAGACTTACCTTCCCTGACCAAGCTGGATATTTTGC 2268
Qy      561 ThrGluSerSerHisPheTrpSerAlaSerAlaLleTrpGlyLysSerAlaSerTrpSer 580
Db      2269 ACAGAGTCTTCACTTCACTTCTCTCTGCTGACCATCTCAAGAGGAGAGCTGAGTTACTCT 2328
Qy      581 AlaTrpSerCysSerGlnLeuProThrCysGlyAspGlnValTrpSerValArgArgArg 600
Db      2329 GCCTACAGCTGCAGCAGAGCTGCCACTTGCAGAGCAAGCTTATTTGTGCGAGCGG 2388
Qy      601 GlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSerProPheGlu 620
Db      2389 CAGAGCCAGAGTGAAGAGCTGACTCGCGGAGCTGCATGAAGAGAGAGAGAGAGAGAGAG 2448
Qy      621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyLysLysLysLysLysLysLys 640
Db      2449 AAGCAGTTTAAACGAGAGAGCTGCCAATTTGGAAGAGCAATGCTCAGAGAAC 2508
Qy      641 ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db      2509 AGCTCAGCGGAAGAGCTGGGGAAGAGTGGGAGTCAAGTTTCTTGGGCAAGATGAA 2568
Qy      661 IleLleGluValSer 665
Db      2569 ATCATTTGAGGCTCTCC 2583

```

RESULT 13
 ID ABV26680 standard; cDNA, 5145 BP.
 AC ABV26680;
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 26671.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-018319P.
 PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI, 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 5388-5389; 11750DP; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX
SO Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other:

Alignment Scores:
Pred. No.: 2.03e-234 Length: 5145
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: 5 Gaps: 0

US-10-029-345A-109 (1-665) x ABV26680 (1-5145)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 589 ATGGCCCATAGATGATTTGAACTCAATGTTACTGAGAGTTGGTGGCTCGCTGGAA 648
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 649 AGTGGAAAGGAAAGAGTGGCTGATTTGATGACCGGCATTTTGGAAATCAATACATCC 708
QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleLeuLeuLeuLeuLeuLeuLeuLeu 60
Db 709 CACATTTTGGAAACCATTAATCAACTGCTCCAAAGTTATGAAAGCAAGTTGCAACAG 768
QY 61 AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Db 769 GAAAGAGTGTAAATTAACAGAGCTATCCAGCATTCACGAAACATTAAGTTGACATTGAT 828
QY 81 CysSerGlnIleValValIleAspGlnSerSerGlnIleAspValAlaSerLeuSerSer 100
Db 829 TGCAGTCAGAAAGTTGATTTAGATCAAGATCCCAAGATTTGCTCTCTCTTCA 888
QY 101 AspCysPheLeuThrValLeuLeuGlyIleLeuGluIleSerPheAsnSerValIleLeu 120
Db 889 GACTGTTTCTCACTGTAATTTGGGTAACTGGAGAAAGACTTCAACTCTGTTCACTG 948
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyIleSer 140
Db 949 CTGGCAGGTGGTTTGTGAGTCTCTCTGTTTCTGCTGCTGCTGAGGAAATTC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 1009 ACTCTAGTCCCTACCTGATTTCTGAGCTTGTCTTACTGTTGCAACATTTGGGCCAAC 1068

QY 161 ArgIleLeuProAsnLeuThrLeuGlyCysGlnArgAspValLeuAsnLeuGluLeu 180
Db 1069 CGAATCTTCCCAATCTTTATCTTGGCTGCAGCGAGATGCTTCAACAAAGAGCTGATG 1128
QY 181 GlnGlnAsnGlyIleGlyThrValLeuAsnAlaSerThrCysAspProLysProAspPhe 200
Db 1129 CAGCAGAAATGGATTTGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 1188
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
Db 1189 ATCCCGAGCTCATTTCTGCGCTGCTGCTGTAAGACAGCTTTGTGAGAAAATTTTG 1248
QY 221 ProThrLeuAspLysSerValAspPheIleGluValAlaValAsnGlyCysVal 240
Db 1249 CCGTGTGGACAAATCAGATGATTCATTGAAAGCAAAACCTTCATGATGATGTGT 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaThrIleMet 260
Db 1309 CTAGTCACCTGTTTACTGTTGGATCTCCGCTCCGCAACATCGCTTACATCATG 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaIleArgPheValIleGluLysArgProThr 280
Db 1369 AAGAGATGGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAGAACTACT 1428
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspThrGluLysIleLysAsn 300
Db 1429 ATATCTCCAAACTCAATTTTCTGGGCAAACTCTGACATGAGAAAGATTAAAGAC 1488
QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuIleLeuLeuHisIleLeuGluLysProAsn 320
Db 1489 CAGACTGGAGCATCAGGCGCAAGAGCAATCAACTGCTGCACTGAGAAAGCCAAAT 1548
QY 321 GluProValProAlaValSerGluGlyGlnLysSerGluThrProLysSerProPro 340
Db 1549 GAACGTGCTCCCTGCTGCTCAGAGGGGTGACAGAAAGCAAGACGCCCTCACTCAACC 1608
QY 341 CysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSerVal 360
Db 1609 TGTGCGACTGTGCTACTCAGAGGAGCAGCAAGAGCCGTGATCCGCGCAGGTG 1668
QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
Db 1669 CCCAGCTGCCAGCTGCAAGCCGCTGCTTTAGAGACAGCCCGCTGTGACAGGCGCTC 1728
QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
Db 1729 AGTGGGCTGCACTGTCCGAGACAGGCTGAAAGACAGCAATAGCTCAAGCGTTCCTTC 1788
QY 401 SerLeuAspIleLysSerValSerThrIleSerAlaSerMetAlaIleSerLeuHisIleGlyPhe 420
Db 1789 TCTCTGATATCAAAATCAGTTTCATATTCAGCCAGATGGCACATCTTACATGGCTTC 1848
QY 421 SerSerSerGlnAspAlaLeuGluIleValIleValProSerThrThrLeuAspGlyThrAsn 440
Db 1849 TCTCTCATCAGAAAGTCTTTGAAATCTACAAACCTTCACTGATGGAGGACCAAC 1908
QY 441 LysLeuCysGlnPheSerProValGlnLeuSerGluGlnThrProGluThrSerPro 460
Db 1909 AAGCTATGCAAGTTCCTCCCTGTTCAAGAACTATCCGAGAGACTCCCGAAACCAAGTCT 1968
QY 461 AspLysGluGluAlaSerIleProLysIleLeuGlnThrAlaArgProSerAspSerGln 480
Db 1969 GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGCAACCGCCAGGCTTCAAGACAGCAG 2028
QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db 2029 AGCAAGCATTTGCTATTCGCTCAGAACCAAGAGAGTGGACCGCCCAAGAGTCCCTTTTA 2088
QY 501 SerProLeuHisArgSerGlySerValGluAspAsnThrHisThrSerPheLeuPheGly 520
Db 2089 TCTCCACTGATGAGAGTGGAGAGCTGAGAGCAATTAACCAACCAAGCTTCTTTTGGGC 2148

QY 521 LeuSerTherSergIngluInleuThyrLysSerAlaGlyLeuLysGlyTrpHis 540
Db 2149 CTTTCCACGACGACGACGACCTCAGAACTGCTGGCTGGCTTAAAGGCTGGCAC 2208
QY 541 SerAspIleuAlaProGlnThrSerThrProSerLeuTherSerTrpTyPheAla 560
Db 2209 TCGGAATATCTGGCCCCCGACGACCTCTACCCCTTCCTGACGACGAGCTGGATTTC 2268
QY 561 ThrGluSerSerHisPheTySerAlaSerAlaIleTyArgLysSerAlaSerTySer 580
Db 2269 ACAGAGCTCTCAGACTTACTCTGCTGACCATCTAGGAGGAGGACGACGACTTACTCT 2328
QY 581 AlaTySerCysSerClnLeuProThrCysGlyAspGlnValTySerValArgArg 600
Db 2329 GCTTACGCTGACGCTGACGCTGCCACTTGGGAGACCAAGCTATTCTGTCGGAGCGG 2388
QY 601 GlnTyPProSerAspArgAlaAspSerArgArgSerTrpHisGlnLysSerProPheGln 620
Db 2389 CAGAAAGCCAAAGTGAAGAAGCTGACCTGCGCGGAGCTGGCATGAAAGAGACCCCTTGA 2448
QY 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGluSerIleMetSerGluAan 640
Db 2449 AAGCAGTTTAAACGACGAAGCTGCCAAATGGAATTGAGAGAGCATCTCAGAGAAC 2508
QY 641 ArgSerArgGlnGluLysGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db 2509 AGGTCAAGGAAAGCTGGGGAAAGTGGGAGGAGTCAAGTCTTTCGGGACGACGTGAA 2568
QY 661 IleIleGlnValSer 665
Db 2569 ATCAATGAGTCTTC 2583

RESULT 14
ABV20978
ID ABV20978 standard; cDNA; 5145 BP.
XX AC ABV20978;
XX DT 13-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 20969.
XX KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KM pharmacogenomic marker; gene; ss.
OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
PS Claim 1; Page 3451; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification of its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 2,03e-234 Length: 5145
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
Gaps: 0

US-10-029-345A-109 (1-665) x ABV20978 (1-5145)

QY 1 MetAlaHisGlnMetIleGlyThrGlnIleValIleThrGluArgLeuValAlaLeuLysGlu 20
Db 589 ATGGCCCATGAGATGATGATGAACTCAATTTGTTACTGAGAGGTTGGTCTGCTGTCGAA 648
QY 21 SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGlnTyPheThrSer 40
Db 649 AGTGAACGGAAGAAAGTGCTGCTAATTGATACCGGCAATTTGGATGATCAATACATCC 708
QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGln 60
Db 709 CACATTTGGAAAGCCATTAATATCAATGCTCAAGCTTATGAAAGCAGAGTTGCAACAG 768
QY 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 769 GACAAAGTGTATATACAGAGCTATCCAGCATTCACGAAACATTAAGTTGACATTGAT 828
QY 81 CysSerGlnLysValValValTyPheAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 829 TGCATGACGAAGGTTGCTGAGTTCTCTCGTTGTTTCCTGGCCTCTGTGAAGAAATCC 888
QY 101 AspCysPheLeuThrValLeuLeuGlnLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db 889 GACTGTTTCTCAGCTGATCTTCTGGTAACTGAGAAAGCTTCAACTCTGTTCACTG 948
QY 121 LeuAlaGlyTyPheAlaGluPheSerArgCysPheProGlyLeuCysGlnGlyLysSer 140
Db 949 CTTCAGGTGGGTTGCTGAGTTCTCTCGTTGTTTCCTGGCCTCTGTGAAGAAATCC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 1009 ACTTACTTCTTACCTGACATTTCTGAGCTTGTACTGTTGCCAACAATTTGGCCAACT 1068
QY 161 ArgIleLeuProAsnLeuTyPheGlyCysGlnArgAspValLeuAsnLysGlnLeuIle 180
Db 1069 CGAATCTTCCCAATCTTATCTTGGCTGCCAGCAGAGTCTCTCAACAAGAGCTGATG 1128
QY 181 GlnGlnAsnGlyIleGlyTyPheValLeuAsnAlaSerTyPheCysProLysProAspPhe 200
Db 1129 CAGCAGAAAGGATTTGTTATGTTAAATGCCACCAATACCTGTCCAAAGCTGACTTT 1188
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
Db 1189 ATCCCGAGATCATTTCTGCTGCTGCTGATGATGACACTTTTGTGAAAAATTTTG 1248
QY 221 ProTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
Db 1249 CCGTGTGGAACAAATCAGTAAATTTCAATGAGAAAGCAAGCTCCCAATGATGTGTT 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyPheIleMet 260

```

Db      1309 CTAGTGACGTTTAAAGTGGAGTCTCCGCTCCGACCACTCCGCTACATCATG 1368
Qy      261  LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleuGlySerProThr 280
Db      1369 AAGAAGATGACATGCTCTTAAAGACTTAAAGATTTTGTGAAGAAAAAAGACCTACT 1428
Qy      281  IleSerProAsnMetPheLeuGlyGlnLeuMetAspTyrGlnValSerIleValAsn 300
Db      1429 ATATCTCAAACTTCAATTTTCTGGCCAACTCTGAGCATATGAGAAAGATTAAAGAAC 1488
Qy      301  GlnThrGlyAlaSerGlyProLysSerLeuLeuIleuGlySerProAsn 320
Db      1489 CAGACTGGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTCGAGAAAGCCAAAT 1548
Qy      321  GluProValProAlaValSerGlyGlyGlnLysSerGluThrProLeuSerProPro 340
Db      1549 GAACCTGTCCCTCTGTCTCAGAGGGTGGACAGAAAGCGAGCGCCCTCAGTCCACCC 1608
Qy      341  CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
Db      1609 TGTGCCGACTCTCTACTCAGAGGAGCAGAGCAAAAGCCGTCGATCCGCGCAGCGTG 1668
Qy      361  ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
Db      1669 CCAGCGTCCGCGCAGCGCGGCGAGCGCTGTTAGAGACAGCCGCTGTGACAGCGCTC 1728
Qy      381  SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
Db      1729 AGTGGGCTGCACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGTCAAGCGTCTTC 1788
Qy      401  SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
Db      1789 TCTCTGATATATCAATCATATTCATATTCAGCCAGCAATGACATCTTACATGGCTTC 1848
Qy      421  SerSerSerGlnAspAlaLeuGlnTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
Db      1849 TCTCTCATCAAAAGATGCTTGGAAATCAACAACCTTCCACTCTGATGGAGCAAC 1908
Qy      441  LysLeuGlnGlnPheSerProValGlnGlnLeuSerGluGlnThrProGluThrSerPro 460
Db      1909 AAGCTATGCGAGTCTCCCTGTTCAAGAACTATCGAGACAGACTCCGAAACAGAGCTC 1968
Qy      461  AspLysGlnGlnAlaSerIleProLysLeuGlnThrAlaArgProSerAspSerGln 480
Db      1969 GATAGAGGAGAACCCAGCATCCCAAGAACCTGAGCGCCAGGCTTACAGACAGCAG 2028
Qy      481  SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db      2029 AGCAAGCGATTCATTCGTCAGAACCAAGCAAGTGGACCGCCCAAGAGTCCCTTTTA 2088
Qy      501  SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
Db      2089 TCTCCACTGCATGGAAGTGGAGGTGGAGACAAATTACCAACACAGCTTCTTTGGCG 2148
Qy      521  LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyThrHis 540
Db      2149 CTTTCCACCGACGACGACCTCAACAACTGCTGGCCCTGAGGCTTAAAGGCTGGCAC 2208
Qy      541  SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
Db      2209 TCGGATATCTTGGCCCCCGAACCTCTACCCCTTCCCTGACAGACACTGATTTTGGC 2268
Qy      561  ThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
Db      2269 ACAGAGTCTCACACTTCTACTGCTGCTCAGCCATCTACGAGAGGAGTGCAGTACTCT 2328
Qy      581  AlaTyrSerGlySerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
Db      2329 GCGTACAGCTGACGACGACGCTGCGCACTTGGGAGAACCAATCTTAATCTTGGCCAGCGG 2388
Qy      601  GlnLysProSerAspArgAlaAspSerArgSerTrpHisGlnGluSerProPheGln 620

```

```

Db      2389 CAGAACCAAGTACAGACACTGACTCGCGCGAGCTGGCATGAAAGACCCCTTTGAA 2448
Qy      621  LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGlnSerIleMetSerGluAsn 640
Db      2449 AAGCAGTTTAAAGCAGAAAGCTGCGCAATGAAATTTGGAGAGGACATCACTAGAGAAC 2508
Qy      641  ArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
Db      2509 AGGTCAAGGAGAGAGCTGGGGAAGTGGCAGTCACTTACGCTTTTGGGACACATGAA 2568
Qy      661  IleIleGluValSer 665
Db      2569 ATCATGAGGTCTCC 2583

RESULT 15
ABV21092
ID      ABV21092 standard; cDNA; 5145 BP.
XX      XX
AC      ABV21092;
XX      XX
DT      13-SEP-2002 (first entry)
XX      XX
DE      Human prostate expression marker cDNA 21083.
XX      XX
KW      Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
        pharmacogenomic marker; gene; ss.
XX      XX
OS      Homo sapiens.
XX      XX
PN      WO200160860-A2.
XX      XX
PD      23-ANG-2001.
XX      XX
PF      20-FEB-2001; 2001WO-US005171.
XX      XX
PR      17-FEB-2000; 2000US-0183119P.
PR      16-MAR-2000; 2000US-0189862P.
PR      25-MAY-2000; 2000US-0207454P.
PR      09-JUN-2000; 2000US-0211314P.
PR      18-JUL-2000; 2000US-0219007P.
PR      13-DEC-2000; 2000US-0255281P.
XX      XX
PA      (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX      XX
PI      Schlegel R, Endege WO, Monahan JR;
XX      XX
DR      WPI; 2001-662795/76.
XX      XX
PT      Novel isolated nucleic acid molecule associated with cancerous state of
PT      prostate cells and correlating with presence of prostate cancer, useful
PT      for detecting presence of prostate cancer, stage of prostate cancer.
XX      XX
PS      Claim 1; Page 3485; 11750bp; English.
XX      XX
CC      The invention relates to an isolated nucleic acid molecule (I) comprising
CC      a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC      specification or its complement. (I) is useful for: (a) assessing whether
CC      a patient is afflicted with prostate cancer; (b) monitoring the
CC      progression of prostate cancer in a patient; (c) assessing the efficacy
CC      of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC      the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC      (e) selecting a composition for inhibiting prostate cancer in a patient;
CC      (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC      determining whether prostate cancer has metastasized in a patient; (h)
CC      assessing the aggressiveness or indolence of prostate cancer in a patient
CC      ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX      XX
SQ      Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:      2,03e-234      Length:      5145
Score:          3406.00      Matches:      663
Percent Similarity: 99.85%      Conservative: 1

```

Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 99.65% Indels: 0
 DB: 5 Gaps: 0

US-10-029-345a-109 (1-665) x ABV21092 (1-5145)

QY 1 MetalaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 DB 589 ATGGCCCATGATGATTGGAACCTCAAAATTGTTACTGAGAGTTGGTGGCTCTGGTGA 648
 QY 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 DB 649 AGTGAAGAGAAAGAGCTGCTGTATTTGATAGCCGGCCATTGTGGAAATACAAATACCC 708
 QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerLeuMetLysArgLeuGln 60
 DB 709 CACATTGGAAAGCCATTAAATCACTGCTCCAAAGCTTATGAAAGCGAAGGTGGCAAG 768
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
 DB 769 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCCAGGAAACATTAAGTTGACATTGAT 828
 QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB 829 TGCAGTCAAGAGTTGTGTATTCATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 888
 QY 101 AspCysPheLeuThrValLeuLeuGluLysLeuGluLysSerPheAsnSerValIleLeu 120
 DB 889 GACGTGTTTCACTGATCTTCTGGTAACTCGAAGAAAGCTTCACTGTGTCACCTG 948
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
 DB 949 CTTCAGAGGTGGATTGCTGAGATTCTCTGTTGTTCCCTGGCCCTCTGGAAGAAATCC 1008
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 1009 ACTCAAGTCCCTACCTGCAATTCCTCAAGCTTCACTGTTCCCAATTGGGCCCAAC 1068
 QY 161 ArgIleLeuProAsnLeuTyrLeuGluCysGlnArgAspValLeuAsnLysGluLeuIle 180
 DB 1069 CGAATTTCTCCCAATCTTATCTTGGCTGCGAGAGATGTCCTCAACAGAGAGCTGATG 1128
 QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
 DB 1129 CAGCAGAAATGGATGTTATGTTAATGCACAGCAATACCTGTCCAAAGCCTGACTTT 1188
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
 DB 1189 ATCCCGAGTCTCATTTCTCGGTGCTGTGAATGACAGCTTTTGAGAGAAATTTTG 1248
 QY 221 ProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
 DB 1249 CCGTGGTTGGCAAAATCATGATTTCTATTGAGAAAGCAAAAGCTCCAAATGATGTGTT 1308
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleLeu 260
 DB 1309 CTAATGCACTGTTAGCTGGAGATCTCCCGCTCCGCAACATCGCTATGCTATCATCATG 1368
 QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
 DB 1369 AAGGAGATGACATGCTTTAGATGAAAGCTTACAGATTTTGAAAGAAAGAAAGCTTACT 1428
 QY 281 IleSerProAsnPheAsnPheLeuGluValLeuLeuAspTyrGluLysLysIleLysAsn 300
 DB 1429 ATATCTCCAAATCTCAATTTTCTGGGCCAACTCTCGAAGTATGAGAAAGATTAAGAAC 1488
 QY 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGluLysProAsn 320
 DB 1489 CAGACTGGAGCATGAGGCGCAAGAGCAAACTCAAGCTGTCGACCTGGAGAAACCAAT 1548
 QY 321 GluProValProAlaValSerGluGlyValGlnLysSerGluThrProLeuSerProPro 340
 DB 1549 GAACCTGTCTCTGTCTCAAGAGGTGGACAGAAAGCGAGACGCCCTTCAATGCCACC 1608

QY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGluArgProValHisProAlaSerVal 360
 DB 1609 TGTGCCACTCTGTACTCTAGAGGCGAGAGCAAAAGCCCTGTGATCCCGCCAGCGTG 1668
 QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
 DB 1669 CCCAGCTGCCAGCGCGAGCGCTGCTGTAGAGACAGCCGCTGTATGAGGCGCTC 1728
 QY 381 SerGlyLeuHisIleuSerAlaAspArgLeuGluAspSerLysLysLeuLysArgSerPhe 400
 DB 1729 AGTGGCTGACCTGTCCGAGACAGGCTGAAAGACAAATAGCTCAAGCTTCTCTTC 1788
 QY 401 SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 DB 1789 TCTCTGATATCAAAATAGATTCTATTAATCCAGACATGACATCTTATCATGTGCTTC 1848
 QY 421 SerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
 DB 1849 TCTCATCAGAAAGATGCTTTGGAATACTCAAACTTCCACTGCTGATGGGACCAAC 1908
 QY 441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
 DB 1909 AAGCTATGCCAGATTCTCCCTGTTCAGAGACATTCGAGAGACTCCCGAAACCAATCT 1968
 QY 461 AspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 DB 1969 GATPAGAGAGAACCCAGCATCCCAAGAACTCCAGACCCCGCTTACAGACAGCAG 2028
 QY 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 DB 2029 AGCAAGAGATGATTCGTGCAGAACAGACAGAGTGGCACCCCGCAAGAGTCCCTTTTA 2088
 QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
 DB 2089 TCTCATGCAATGAGAGTGGAGCGTGGAGAGCAATTAACACACAGACTTCTTTTGCC 2148
 QY 521 LeuSerThrSerGlnGlnHisIleuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheHis 540
 DB 2149 CTTTCCACAGCCAGACAGACACTCAGAAATCTGCTGGCCCTGGGCTTAAGGCTGGCAC 2208
 QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
 DB 2209 TCGGATATCTTGGCCCCCGACCTTACCTCCCTGACACAGAGCTGTATTTTGCC 2268
 QY 561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSer 580
 DB 2269 ACGAGTCTCAACAATTCTACTGCTGCTCAGCAATCTACGAGGCAATGCCAGATTACTCT 2328
 QY 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
 DB 2329 GCTTACAGCTGACGACGACGCTGCCACTTGGCGAGACCAAGTCTAATCTTGCGCAGGCGG 2388
 QY 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrPheGluLeuSerProPheGlu 620
 DB 2389 CAGAAGCAAGTACAGAGCTGACTCGCGCGAGCTGGCATTAAGAGAGCCCTTTTGA 2448
 QY 621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyValLeuSerIleMetSerGluAsn 640
 DB 2449 AAGCAGTTTAAAGCAGAACTCCCAATGCAATTTGGAGAGCATCATGTCAAGAAC 2508
 QY 641 ArgSerArgGluLeuGluLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 DB 2509 AGGTCAAGGAAAGAGCTGGGAAAGTGGGAGTCAAGTCTTCTTCCGCAAGATGAA 2568
 QY 661 IleIleGluValSer 665
 DB 2569 ATCATTGAGGCTTCC 2583

Search completed: June 21, 2004, 21:32:57
 Job time : 776.933 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:10:27 ; Search time 74.2709 Seconds

(without alignments)
2529.847 Million cell updates/sec

Title: US-10-029-345A-109

Sequence: 1 MAHEMIGQIVTERLVALL.....LGKVGSSFSFGSMELIEVS 665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:***
2: geneseqp1980s:***
3: geneseqp2000s:***
4: geneseqp2001s:***
5: geneseqp2002s:***
6: geneseqp2003as:***
7: geneseqp2003bs:***
8: geneseqp2004s:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3418	100.0	665	5	ABR52381 Protein r
2	3418	100.0	665	5	ABR52407 Protein r
3	3406	99.6	665	4	AAE04834 Human SGP
4	3406	99.6	665	4	AAU09016 Human dua
5	3406	99.6	665	5	AAU79156 Human dua
6	3406	99.6	665	5	AAU09946 Protein s
7	3406	99.6	665	5	AAU75789 Human pro
8	3406	99.6	665	5	ABR97946 Human pro
9	3406	99.6	665	5	AAU79929 Human dua
10	3406	99.6	665	5	ABR97291 Novel hum
11	3399	99.4	665	6	ADA54744 Human pro
12	3398	99.4	665	5	AAU79161 Human dua
13	3396	99.4	665	5	AAU79162 Human dua
14	3388	99.1	665	5	ABR52352 Protein r
15	3368.5	98.6	664	5	ABR52424 Protein r
16	3057.5	89.5	660	5	ABR52385 Protein r
17	3030.5	88.7	666	4	AAAB20325 Human pro
18	2930	85.7	672	4	AAAB25744 Human pro
19	2500	73.1	517	5	AAU79159 Human dua
20	1542	45.1	302	5	ABR52425 Protein r
21	1326	38.8	625	5	ABR52382 Protein r
22	1326	38.8	625	5	ABR52350 Protein r
23	1326	38.8	625	6	ABG73440 Human dua
24	1302	38.1	663	5	ABR52351 Protein r
25	1297	37.9	663	2	AAW29150 Dual-spec

26	977.5	28.6	579	7	ABR52381	ABR52381 standard; protein, 665 AA.
27	742.5	21.7	375	4	ABR52381	ABR52381 standard; protein, 665 AA.
28	741	21.7	140	5	ABR52404	ABR52404 Peptide r
29	716	20.9	140	5	ABR52405	ABR52405 Peptide r
30	654.5	19.1	170	4	ABR52436	ABR52436 Human MAP
31	654.5	19.1	170	4	AAE06780	AAE06780 Human dua
32	476	13.9	482	4	AAAB3225	AAAB3225 Human pro
33	471.5	13.8	381	5	ABR52383	ABR52383 Protein r
34	471.5	13.8	381	6	ABR52383	ABR52383 Protein r
35	469	13.7	444	5	AAO20515	AAO20515 Protein o
36	469	13.7	482	3	AAAB18655	AAAB18655 A human r
37	469	13.7	482	3	AAAB29641	AAAB29641 Human dua
38	469	13.7	482	5	ABR52384	ABR52384 Protein r
39	469	13.7	482	5	ABR52384	ABR52384 Protein r
40	466.5	13.6	381	6	ABR57087	ABR57087 Mouse MKP
41	465.5	13.6	381	7	ADD48300	ADD48300 Rat Prote
42	465.5	13.6	381	7	ADG62625	ADG62625 Rat Prote
43	453	13.3	394	4	AAAB76875	AAAB76875 Human lun
44	453	13.3	394	5	AAU85530	AAU85530 STY8 lung
45	453	13.3	394	6	ABU69502	ABU69502 Humna STY

ALIGNMENTS

RESULT 1
ID ABR52381 standard; protein, 665 AA.
AC ABR52381;
XX
XX
DT 19-JUN-2003 (first entry)
XX
XX
DE Protein relating to the invention SEQ ID NO: 109.
XX
XX
KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;
KW antiproliferative; cardiac; cytosolic; gene therapy; liver disease;
KW proliferative disorder; renal failure; cardiovascular disorder;
KW immunological disorder; arthritis; psoriasis; congenital heart defect;
KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
XX
OS Homo sapiens.
XX
XX
PN MO200257460-A2.
XX
XX
PD 25-JUL-2002.
XX
XX
PF 20-DEC-2001; 2001MO-US050459.
XX
XX
PR 20-DEC-2000; 2000US-0256688P.
PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX
PI Jackson DG, Feder J, Nelson T, Mintier G, Ramamathan C, Lee L;
PI Stemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
PI Kyurek S, Mcatee P, Suchard S, Banas D;
XX
XX
DR WPI: 2002-599721/64.
DR N-PSDB; ACC60559.
XX
XX
PT Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
XX
PS Claim 5; Fig 12; 801p; English.
XX
XX
CC The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of

CC the invention has antiproliferative, hepatotropic, nephrotropic, antiarthritis, antipsoriatic, cardiant, and cytostatic activity. The polynucleotide may have a use in gene therapy. A polynucleotide or polypeptide of the invention is useful for preventing, treating or ameliorating a medical condition, e.g. a proliferative disorder. They are also useful for treating e.g. liver disease, renal failure, immunological disorders including arthritis and psoriasis, cardiovascular disorders such as congenital heart defects and congestive heart failure, and cancer. A method of the invention is useful for diagnosing a pathological condition or susceptibility to a condition in a subject. The present sequence is used in the exemplification of the invention.

Sequence 665 AA:

Query Match 100.0%; Score 3418; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 3.1e-257;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAHEMIGIYTERLVALLSGTEKVLIDSRPVEVNTSHLEAININSGKLMKRLQ 60
DB 1 MAHEMIGIYTERLVALLSGTEKVLIDSRPVEVNTSHLEAININSGKLMKRLQ 60
QY 61 DKVLITELIHSAAKHVIDCSQKVVYVYDOSQDVASLSDDCFVYLGLKESFNSVHL 120
DB 61 DKVLITELIHSAAKHVIDCSQKVVYVYDOSQDVASLSDDCFVYLGLKESFNSVHL 120
QY 121 LAGFAEFSRCFPGICGKSTLVPTCISQCLPVANIGPRLIPNLYLGGQDVANKELI 180
DB 121 LAGFAEFSRCFPGICGKSTLVPTCISQCLPVANIGPRLIPNLYLGGQDVANKELI 180
QY 181 QONGIGVYLNASYCPRPDFIPESHFLRPVNSFCEKILPWLDSKVDPIERAKASNGCV 240
DB 181 QONGIGVYLNASYCPRPDFIPESHFLRPVNSFCEKILPWLDSKVDPIERAKASNGCV 240
QY 241 LVHCLAGISRSATTAIYIMKRMDSIDEAYRPFVKEKRTISPNNFLGQLIDYEKKIKN 300
DB 241 LVHCLAGISRSATTAIYIMKRMDSIDEAYRPFVKEKRTISPNNFLGQLIDYEKKIKN 300
QY 301 QTGASGPKSLKULHLEKNEPVPVASEGQSETPPLSPCADSATSTSAAGQRPVHPASV 360
DB 301 QTGASGPKSLKULHLEKNEPVPVASEGQSETPPLSPCADSATSTSAAGQRPVHPASV 360
QY 361 PSVPSVQPSLIDSPVQALSGHLISADRLSDNKLKRSFSLDIKSVYSASMAASLHGF 420
DB 361 PSVPSVQPSLIDSPVQALSGHLISADRLSDNKLKRSFSLDIKSVYSASMAASLHGF 420
QY 421 SSSSEDALEYKRPSTLIDGTNKLCOFSPVOELSEOTPEPSPKXEAASIPKLTQTAAPSDQ 480
DB 421 SSSSEDALEYKRPSTLIDGTNKLCOFSPVOELSEOTPEPSPKXEAASIPKLTQTAAPSDQ 480
QY 481 SKRLHSVTSSTSGTAQRSLSPLRSGSVEDNHTSFLFGLSTISQOHLTKSAGLKGWH 540
DB 481 SKRLHSVTSSTSGTAQRSLSPLRSGSVEDNHTSFLFGLSTISQOHLTKSAGLKGWH 540
QY 541 SDIILAPQTSPTSLTSSWYFATESHFPASASAYGSGASYSAGSCQLPTCGQVYSVRRR 600
DB 541 SDIILAPQTSPTSLTSSWYFATESHFPASASAYGSGASYSAGSCQLPTCGQVYSVRRR 600
QY 601 QKPSDRADSRSMHEESPFEKOFRRSCQWEPGESIMSENRSREELGKYGSSSFGSME 660
DB 601 QKPSDRADSRSMHEESPFEKOFRRSCQWEPGESIMSENRSREELGKYGSSSFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

```

RESULT 2
ABR52407 standard; protein; 665 AA.
AC ABR52407;
XX

DT 19-JUN-2003 (first entry)
XX
DE Protein relating to the invention SEQ ID NO: 148.
XX
KW antiproliferative; hepatotropic; nephrotropic; antiarthritis;
KW antipsoriatic; cardiant; cytostatic; gene therapy; liver disease;
KW proliferative disorder; renal failure; cardiovascular disorders;
KW immunological disorder; arthritis; psoriasis; congenital heart defect;
KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
OS Homo sapiens.
XX
PN WO200257460-A2.
XX
PD 25-JUN-2002.
XX
PF 20-DEC-2001; 2001WO-US050459.
XX
PR 20-DEC-2000; 2000US-0256686P.
XX
PR 30-MAR-2001; 2001US-0280186P.
XX
PR 01-MAY-2001; 2001US-0287735P.
XX
PR 05-JUN-2001; 2001US-0295648P.
XX
PR 25-JUN-2001; 2001US-0300465P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
PI Krystek S, Mcatee P, Suchard S, Banaas D;
XX
DR WPI; 2002-599721/64.
XX
DR N-P-SDB; ACC60572.

PT Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.

PS Disclosure; Fig 19; 801pp; English.

XX The invention relates to a novel isolated nucleic acid comprising a
XX polynucleotide having a nucleotide sequence selected from 40
XX polynucleotides fully defined in the specification. The polynucleotide of
XX the invention has antiproliferative, hepatotropic, nephrotropic,
XX antiarthritis, antipsoriatic, cardiant, and cytostatic activity. The
XX polynucleotide may have a use in gene therapy. A polynucleotide or
XX polypeptide of the invention is useful for preventing, treating or
XX ameliorating a medical condition, e.g. a proliferative disorder. They are
XX also useful for treating e.g. liver disease, renal failure, immunological
XX disorders including arthritis and psoriasis, cardiovascular disorders
XX such as congenital heart defects and congestive heart failure, and
XX cancer. A method of the invention is useful for diagnosing a pathological
XX condition or susceptibility to a condition in a subject. The present
XX sequence is used in the exemplification of the invention

Sequence 665 AA:

Query Match 100.0%; Score 3418; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 3.1e-257;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAHEMIGIYTERLVALLSGTEKVLIDSRPVEVNTSHLEAININSGKLMKRLQ 60
DB 1 MAHEMIGIYTERLVALLSGTEKVLIDSRPVEVNTSHLEAININSGKLMKRLQ 60
QY 61 DKVLITELIHSAAKHVIDCSQKVVYVYDOSQDVASLSDDCFVYLGLKESFNSVHL 120
DB 61 DKVLITELIHSAAKHVIDCSQKVVYVYDOSQDVASLSDDCFVYLGLKESFNSVHL 120
QY 121 LAGFAEFSRCFPGICGKSTLVPTCISQCLPVANIGPRLIPNLYLGGQDVANKELI 180
DB 121 LAGFAEFSRCFPGICGKSTLVPTCISQCLPVANIGPRLIPNLYLGGQDVANKELI 180
QY 181 QONGIGVYLNASYCPRPDFIPESHFLRPVNSFCEKILPWLDSKVDPIERAKASNGCV 240

```

```

Db      181 QQNGIGVNLASTYCPKDPDI PESHFLRVVNDSECKILPMDLXSDPIEKKASNGCV 240
Qy      241 LVHCLAGISRSATTAIAIYIMKRDMSIDEAYRVEKERPTISPNFNLGOLLDEKKIKN 300
Db      241 LVHCLAGISRSATTAIAIYIMKRDMSIDEAYRVEKERPTISPNFNLGOLLDEKKIKN 300
Qy      301 QTGASGPKSKLKLHLKEKPEVPVAVSEGQKSETPSPCADSATSEAGRPVHPASV 360
Db      301 QTGASGPKSKLKLHLKEKPEVPVAVSEGQKSETPSPCADSATSEAGRPVHPASV 360
Qy      361 PSVPSVQPSLLEDSPLVQALSGHLASDRLEDNKLKRSFLDIKSVSASMAASLHGF 420
Db      361 PSVPSVQPSLLEDSPLVQALSGHLASDRLEDNKLKRSFLDIKSVSASMAASLHGF 420
Qy      421 SSSBDALYYKPESTTLDGTNKLCOFSPVQELSEQTPEPSDKEEASIPKLTQTPARPSDSQ 480
Db      421 SSSBDALYYKPESTTLDGTNKLCOFSPVQELSEQTPEPSDKEEASIPKLTQTPARPSDSQ 480
Qy      481 SKRLHSVRTSSGCTAORSLSPLRSGSVEDNHTSLRGLSTNSQOHLTKSAGLGLKGMH 540
Db      481 SKRLHSVRTSSGCTAORSLSPLRSGSVEDNHTSLRGLSTNSQOHLTKSAGLGLKGMH 540
Qy      541 SDILAPQTSPTSLTSSWYFATESHFFASAIYGSASVAVSCQPLTCGDQVYVRRR 600
Db      541 SDILAPQTSPTSLTSSWYFATESHFFASAIYGSASVAVSCQPLTCGDQVYVRRR 600
Qy      601 QKPSDRADRSRRSWHEESPFEKQFRRSCQMEFGESIMSENRREELKGVSSQSPSGSME 660
Db      601 QKPSDRADRSRRSWHEESPFEKQFRRSCQMEFGESIMSENRREELKGVSSQSPSGSME 660
Qy      661 IIEVS 665
Db      661 IIEVS 665

```

RESULT 3

AAE04834
ID AAE04834 standard; protein: 665 AA.

AC AAE04834;
XX
DT 10-SEP-2001 (first entry)
XX

DE Human SGP002 phosphatase polypeptide.

XX Human; SGP002 phosphatase polypeptide; phosphatase-related disease;
KM Immune-related disorder; ocular disease; organ transplant rejection;
KM Infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KM Metabolic disorder; hematopoietic cancer; mood disorder; cardiac;
KM Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KM Cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KM Attention disorder; cognition disorder; psychotic disorder; cytostatic;
KM Neurological disorder; vitruclide; noctropic; cerebroprotective; therapy;
KM Neuroprotective; antibacterial; antineoplastic; tranquilizer; antidiabetic;
KM hypotensive; immunosuppressive; antineoplastic; analgesic; hypertensive;
KM antifungal; dual specificity phosphatase; DSP; Map kinase phosphatase;
KM MKP; mitraline; chromosome 12p11.1-p12.1.
XX

OS Homo sapiens.

XX Key Location/Qualifiers

FT Domain 1..173
FT Domain /label= Catalytic_domain
FT Domain 158..297
FT Domain /label= Phosphatase_domain

XX MO200146394-A2.

XX 28-JUN-2001.

XX 21-DEC-2000; 2000MO-US034736.

XX

```

PR 21-DEC-1999; 99US-0173255P.
PR 28-DEC-1999; 99US-0175766P.
PR 25-JAN-2000; 2000US-0178078P.
PR 31-JAN-2000; 2000US-0179301P.
XX (SUGC-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarshanam S;
PI Hill RJ, Flanagan P;
XX
DR WPI; 2001-418058/44.
DR N-PSDB; AAD09492.
XX
PT Novel phosphatase polypeptide useful for treating cancers, immune-related
PT diseases and disorders, cardiovascular disease, brain or neuronal-
PT associated diseases and metabolic disorders.
XX
PS Claim 7; Fig 2; 186pp; English.
XX

```

The present invention relates to phosphatase polypeptides, nucleotide sequences encoding them, as well as various products and methods useful for the diagnosis and treatment of various phosphatase-related diseases and conditions. Substance that modulates the activity of phosphatase polypeptide is used to treat immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases and metabolic disorders, including cancers of tissue, cancers of haematopoietic origin, diseases of central and peripheral nervous system, Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, viral infections, infections caused by prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, hypertension, psychotic disorders, neurological disorders, dyskinesias and organ transplant rejection. The present amino acid sequence is human SGP002 phosphatase polypeptide. This sequence is classified as dual specificity phosphatase (DSP) and Map kinase phosphatase (MKP). SGP002 gene maps to chromosome position 12p11.1-p12.1

Sequence 665 AA:

Query Match 99.6%; Score 3406; DB 4; Length 665;
Best Local Similarity 99.7%; Pred. No. 2,7e-256;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 MAHEMGTQIVTERLVALLESGETEKLIDSRPFVENYSHILEATININSGKMKRLQ 60
Db      1 MAHEMGTQIVTERLVALLESGETEKLIDSRPFVENYSHILEATININSGKMKRLQ 60
Qy      61 DKVLITELQHSARKHVDIDCSQKVYVDQSODVASLSDCFLTVLLGLKESFNSVHL 120
Db      61 DKVLITELQHSARKHVDIDCSQKVYVDQSODVASLSDCFLTVLLGLKESFNSVHL 120
Qy      121 LAGFAEFRCFPGLCGKSTLVPTCISQCLPVANIGPFRILPNLYLGGORDVANKEL 180
Db      121 LAGFAEFRCFPGLCGKSTLVPTCISQCLPVANIGPFRILPNLYLGGORDVANKEL 180
Qy      181 QQNGIGVNLASTYCPKDPDI PESHFLRVVNDSECKILPMDLXSDPIEKKASNGCV 240
Db      181 QQNGIGVNLASTYCPKDPDI PESHFLRVVNDSECKILPMDLXSDPIEKKASNGCV 240
Qy      241 LVHCLAGISRSATTAIAIYIMKRDMSIDEAYRVEKERPTISPNFNLGOLLDEKKIKN 300
Db      241 LVHCLAGISRSATTAIAIYIMKRDMSIDEAYRVEKERPTISPNFNLGOLLDEKKIKN 300
Qy      301 QTGASGPKSKLKLHLKEKPEVPVAVSEGQKSETPSPCADSATSEAGRPVHPASV 360
Db      301 QTGASGPKSKLKLHLKEKPEVPVAVSEGQKSETPSPCADSATSEAGRPVHPASV 360
Qy      361 PSVPSVQPSLLEDSPLVQALSGHLASDRLEDNKLKRSFLDIKSVSASMAASLHGF 420
Db      361 PSVPSVQPSLLEDSPLVQALSGHLASDRLEDNKLKRSFLDIKSVSASMAASLHGF 420
Qy      421 SSSBDALYYKPESTTLDGTNKLCOFSPVQELSEQTPEPSDKEEASIPKLTQTPARPSDSQ 480
Db      421 SSSBDALYYKPESTTLDGTNKLCOFSPVQELSEQTPEPSDKEEASIPKLTQTPARPSDSQ 480

```

Db 421 SSSDALFYKPKSTTLIDGNTKLCQFSPVOELSBQTPETSPDKERASIPKLTQTARPSDSQ 480
QY 481 SKRLHSVTRSSSGTQORSLISPLHRSGSVEDNYHTSFLFGISTSQOHLTKSAGLKGWH 540
Db 481 SKRLHSVTRSSSGTQORSLISPLHRSGSVEDNYHTSFLFGISTSQOHLTKSAGLKGWH 540
QY 541 SDIAPQTSPTSLTSSWYFATESSHFYASAIYGGASYSAYSCSOLPTCGDQVYSVRRR 600
Db 541 SDIAPQTSPTSLTSSWYFATESSHFYASAIYGGASYSAYSCSOLPTCGDQVYSVRRR 600
QY 601 QKPSDRADSRSRWHEESPFEKQFKRRSCQMEFGESIMSENRSREELGKVGSSSFGSME 660
Db 601 QKPSDRADSRSRWHEESPFEKQFKRRSCQMEFGESIMSENRSREELGKVGSSSFGSME 660
QY 661 IIEVS 665
Db 661 IIEVS 665

RESULT 4
AAU09016
ID AAU09016 standard; protein; 665 AA.
XX AAU09016;

DT 18-DEC-2001 (first entry)
XX
DE Human dual specificity phosphatase 21117.

XX Human; dual specificity phosphatase 21117; hepatotropic; cytostatic;
KM hematoepietic disorder; autoimmune disorder; diabetes mellitus;
KM rheumatoid arthritis; multiple sclerosis; Crohn's disease;
KM liver disorder; erythroid associated disorder; haemolytic anaemia;
KM cellular proliferative; differentiative disorder; leukaemia;
KM acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer;
XX immunogen.

OS Homo sapiens.

XX Key
FH Location/Qualifiers
FT 11..131

FT Domain
FT /label= Rhodanese-like domain
FT 158..297

FT Domain
FT /label= Catalytic domain
FT /note= "Dual specificity phosphatase catalytic domain"
FT 242..254

FT Active-site
FT /label= Tyrosine_specific_protein_phosphatase_active_site
PN W0200173059-A2.

XX 04-OCT-2001.

XX 23-MAR-2001, 2001WO-US009477.

XX 24-MAR-2000, 2000US-0191858P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Meyers RA;

XX WPI; 2001-611635/70.

XX N-PSDB; AAS14639.

XX New human dual specificity polypeptides and nucleic acids for diagnosis
PT of disease and treatment of e.g. liver disorders.

XX Claim 9; Fig 1; 143pp; English.

XX The invention relates to two novel human dual specificity phosphatases
CC designated 21117 and 38692, the nucleic acids encoding them (including
CC fragments, allelic variants, their complements or nucleic acids that
CC hybridise to them) and antibodies raised against the proteins. The
CC antibody is useful for detecting the presence of the polypeptide, and the

CC nucleic acid fragments are useful for detecting the presence of the
CC nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
CC antisense sequences) are useful for modulating the activity or expression
CC of the polypeptide/nucleic acid, useful for the treatment of e.g.
CC hematoepietic disorders, autoimmune disorders (e.g. diabetes mellitus,
CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
CC listed in the specification) liver disorders, erythroid associated
CC disorders (e.g. haemolytic anaemia) cellular proliferative or
CC differentiative disorders, leukaemias (e.g. acute myeloid leukaemia),
CC carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and
CC 38692 are also useful for modulating the proliferation, survival,
CC migration or differentiation of a 38692 or 21117-expressing cell. The
CC polypeptide and nucleic acids are useful for identifying modulating
CC agents. The present sequence represents the dual specificity phosphatase
XX 21117

SQ Sequence 665 AA;

Query Match 99.6%; Score 3406; DB 4; Length 665;
Best Local Similarity 99.7%; Pred. No. 2..7e-256;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTOIVTEBELVALLESSTGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRRLOQ 60
Db 1 MAHEMIGTOIVTEBELVALLESSTGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIQSHAKHVDIDCSQKVVYVDOSSQDVASLSDDCFLLVLLGKLEKSFNSVHL 120
Db 61 DKVLITELIQSHAKHVDIDCSQKVVYVDOSSQDVASLSDDCFLLVLLGKLEKSFNSVHL 120
QY 121 LAGGFARFRCFPGLCCEKSTLVPTCISQCPVANIPTLILPNLYGCGQDVLANKELI 180
Db 121 LAGGFARFRCFPGLCCEKSTLVPTCISQCPVANIPTLILPNLYGCGQDVLANKELI 180
QY 121 LAGGFARFRCFPGLCCEKSTLVPTCISQCPVANIPTLILPNLYGCGQDVLANKELI 180
Db 121 LAGGFARFRCFPGLCCEKSTLVPTCISQCPVANIPTLILPNLYGCGQDVLANKELI 180
QY 181 QONGIGYVNAASYTCPKDFIPESHFLVPPVNDSCFKILPWLDKSVDFTEKAKASNGCV 240
Db 181 QONGIGYVNAASYTCPKDFIPESHFLVPPVNDSCFKILPWLDKSVDFTEKAKASNGCV 240
QY 181 QONGIGYVNAASYTCPKDFIPESHFLVPPVNDSCFKILPWLDKSVDFTEKAKASNGCV 240
Db 181 QONGIGYVNAASYTCPKDFIPESHFLVPPVNDSCFKILPWLDKSVDFTEKAKASNGCV 240
QY 241 LVHCLAGISRSATYIAIYIMKRMDSIDEAYRFYKERRPTISPFNFLLGQLDYEEKIKN 300
Db 241 LVHCLAGISRSATYIAIYIMKRMDSIDEAYRFYKERRPTISPFNFLLGQLDYEEKIKN 300
QY 301 QTGASGPKSKLKLHLKPNPVPVAVSGQKSTPLSPPCADSATSEAAQRPVHPASV 360
Db 301 QTGASGPKSKLKLHLKPNPVPVAVSGQKSTPLSPPCADSATSEAAQRPVHPASV 360
QY 361 PSVSVQPSLLEDSPVQALSGHLNSADRLSDNKLKRSFSLDINSVSYSAWMAASLHGF 420
Db 361 PSVSVQPSLLEDSPVQALSGHLNSADRLSDNKLKRSFSLDINSVSYSAWMAASLHGF 420
QY 421 SSSDALFYKPKSTTLIDGNTKLCQFSPVOELSBQTPETSPDKERASIPKLTQTARPSDSQ 480
Db 421 SSSDALFYKPKSTTLIDGNTKLCQFSPVOELSBQTPETSPDKERASIPKLTQTARPSDSQ 480
QY 481 SKRLHSVTRSSSGTQORSLISPLHRSGSVEDNYHTSFLFGISTSQOHLTKSAGLKGWH 540
Db 481 SKRLHSVTRSSSGTQORSLISPLHRSGSVEDNYHTSFLFGISTSQOHLTKSAGLKGWH 540
QY 541 SDIAPQTSPTSLTSSWYFATESSHFYASAIYGGASYSAYSCSOLPTCGDQVYSVRRR 600
Db 541 SDIAPQTSPTSLTSSWYFATESSHFYASAIYGGASYSAYSCSOLPTCGDQVYSVRRR 600
QY 601 QKPSDRADSRSRWHEESPFEKQFKRRSCQMEFGESIMSENRSREELGKVGSSSFGSME 660
Db 601 QKPSDRADSRSRWHEESPFEKQFKRRSCQMEFGESIMSENRSREELGKVGSSSFGSME 660
QY 661 IIEVS 665
Db 661 IIEVS 665

RESULT 5
AAU79156

ID AAV79156 standard; protein; 665 AA.
 AC AAV79156;
 DT 02-JUL-2002 (first entry)
 DE Human dual-specificity phosphatase-3 (DSP-16) protein.
 XX
 XX Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KM mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KM cancer; graft-versus-host disease; allergy; metabolic disease;
 KM abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KM cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KM intercellular adhesion; DSP-16 modulator; chromosome 12p.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT Domain 242..251
 FT /label= Active_site_domain
 MO200226997-A2.
 PD 04-APR-2002.
 XX
 XX 25-SEP-2001; 2001MO-US030124.
 PF
 XX
 XX 26-SEP-2000; 2000US-0235487P.
 PR
 XX
 XX (CEPT-) CEPTYR INC.
 PA
 XX
 XX Luche RM, Wei B;
 PI
 XX
 XX WPI; 2002-315802/35.
 DR N-PSDB; ABK47596.
 XX
 XX New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 PS
 XX
 PS Claim 1; Fig 2; 87pp; English.
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity, for modulation of a proliferative response in a cell, survival
 CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present amino acid
 CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 CC protein of the invention. This sequence is encoded by the human DSP-16
 CC gene located on chromosome 12p
 CC
 CC
 SO Sequence 665 AA;
 Query Match 99.6%; Score 3406; DB 5; Length 665;
 Best Local Similarity 99.7%; Pred. No. 2.7e-256;
 Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MAHEMIGTQVTERLVALLSGTEKVLIDSRPVEVNIHSHILKAININSGKMKRRLQ 60
 DB 1 MAHEMIGTQVTERLVALLSGTEKVLIDSRPVEVNIHSHILKAININSGKMKRRLQ 60
 QY 61 DKVLITELIHSAKHKVDICQKRVVYDOSQDVASLSSDCEFLTVLLGLKESFNSVHL 120
 DB 61 DKVLITELIHSAKHKVDICQKRVVYDOSQDVASLSSDCEFLTVLLGLKESFNSVHL 120

QY 121 LAGFAEFSRCFPGLCCEKSTLVPTCISQPCLPVANIIPRIIPNLVLCQDVLANKELI 180
 DB 121 LAGFAEFSRCFPGLCCEKSTLVPTCISQPCLPVANIIPRIIPNLVLCQDVLANKELI 180
 QY 181 QQNGIGVNLNASTYCKPDPFIPESHFLRVVNDSPCEKILPWLKSVDFIEKASNGCV 240
 DB 181 QQNGIGVNLNASTYCKPDPFIPESHFLRVVNDSPCEKILPWLKSVDFIEKASNGCV 240
 QY 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRVYKERRPTISNPNFLGOLLDYKKIKN 300
 DB 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRVYKERRPTISNPNFLGOLLDYKKIKN 300
 QY 301 QTSASGPKSKLKLHLKEMPEVPVAVSEGQKSETPSPCCASATSSEAGORPVHPASV 360
 DB 301 QTSASGPKSKLKLHLKEMPEVPVAVSEGQKSETPSPCCASATSSEAGORPVHPASV 360
 QY 361 PSVPSVQPSLIEDSPVLVALSGHLASDRLEDNKLKRSFLDIKSVYSASMAASLHGF 420
 DB 361 PSVPSVQPSLIEDSPVLVALSGHLASDRLEDNKLKRSFLDIKSVYSASMAASLHGF 420
 QY 421 SSSSEDALLEYKPESTTIDGTNKLCOFSPVQELSQETPETSDEKESASIPKQLQIARPSDSQ 480
 DB 421 SSSSEDALLEYKPESTTIDGTNKLCOFSPVQELSQETPETSDEKESASIPKQLQIARPSDSQ 480
 QY 481 SKRLHSVRTSSSGTAORSLISPLHRSQSVEDNYHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
 DB 481 SKRLHSVRTSSSGTAORSLISPLHRSQSVEDNYHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
 QY 541 SDILAPQTSPLSLTSSWYFATBSSHFSYASAIYGSASYSAYSCSLPTCGDQVYSVRRR 600
 DB 541 SDILAPQTSPLSLTSSWYFATBSSHFSYASAIYGSASYSAYSCSLPTCGDQVYSVRRR 600
 QY 601 QKPSDPAISRHSWHESPPFKQFKRRSCOMEFGESIMSENRREBELGKVGSSFSQSM 660
 DB 601 QKPSDPAISRHSWHESPPFKQFKRRSCOMEFGESIMSENRREBELGKVGSSFSQSM 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665
 RESULT 6
 AAV09946
 ID AAV09946 standard; protein; 665 AA.
 AC AAV09946;
 DT 18-JUN-2002 (first entry)
 DE Protein sequence of human (dual specificity phosphatase) DSP-10.
 XX
 XX Human; dual specificity phosphatase; DSP-10; cancer; epilepsy; stroke;
 KM neuronal degeneration syndrome; Alzheimer's disease; depression;
 KM schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
 KM osteoporosis; diabetes.
 OS Homo sapiens.
 XX
 XX
 PN MO200177340-A1.
 PD 18-OCT-2001.
 XX
 XX 06-APR-2001; 2001MO-EP003966.
 PF
 XX
 XX 10-APR-2000; 2000EP-00107143.
 PR
 XX
 XX (MERE) MERCK PATENT GBM.
 PA
 XX
 XX Duecker K;
 PI
 XX
 XX WPI; 2002-010917/01.
 DR N-PSDB; AAS15768.
 XX

PT Novel dual specificity phosphatase polypeptides useful for treating
PT cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
PT disease, depression, schizophrenia, asthma and immune disorders.
XX
XX Claim 2; Page 37-39; 43pp; English.

CC The present invention relates to a new isolated dual specificity
CC phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
CC sequence that is fully defined in the specification. The invention also
CC provides a sequence encoded by a 3059 nucleotide sequence fully defined
CC in the specification, and a sequence having at least 95 % identity to the
CC polypeptide, or fragments or variants of DUSP-10. The invention is useful
CC for treating cancer e.g. leukaemia, colon carcinoma, lung cancer,
CC prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
CC stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
CC schizophrenia, cardiac myopathies, asthma, immune disorders,
CC osteoporosis, diabetes and diabetes associated diseases. The molecules of
CC the invention are also useful as vaccines for inducing immunological
CC response in a mammal, in disease diagnosis and in assays for screening
CC agonistic or antagonistic compounds. Other uses of the invention include
CC identifying membrane bound or soluble receptors, as a diagnostic reagent,
CC in chromosome localization studies, and as a valuable tool in tissue
CC expression studies. The present sequence represents the dual specificity
CC phosphatase, DUSP-10, protein of the invention

XX Sequence 665 AA;

Query Match 99.6%; Score 3406; DB 5; Length 665;
Best Local Similarity 99.7%; Pred. No. 2.7e-256;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGQIVTERLVALLSGETEKVLLIDSRPVENYSHIIEAININCSKLMKRRLLQ 60
DB 1 MAHEMIGQIVTERLVALLSGETEKVLLIDSRPVENYSHIIEAININCSKLMKRRLLQ 60
QY 61 DKVLTTELIOHSAHAKVDIDCSOKVVYDSSDVASISDCLFVTLGKLEKSNVHL 120
DB 61 DKVLTTELIOHSAHAKVDIDCSOKVVYDSSDVASISDCLFVTLGKLEKSNVHL 120
QY 121 LAGGFAEFRCFPELCEGKSTLVPTCISOPCLPVANTGTRLLPMLVYLGCCRDVLMKELI 180
DB 121 LAGGFAEFRCFPELCEGKSTLVPTCISOPCLPVANTGTRLLPMLVYLGCCRDVLMKELI 180
QY 121 LAGGFAEFRCFPELCEGKSTLVPTCISOPCLPVANTGTRLLPMLVYLGCCRDVLMKELI 180
DB 121 LAGGFAEFRCFPELCEGKSTLVPTCISOPCLPVANTGTRLLPMLVYLGCCRDVLMKELI 180
QY 181 QONGIGVYVMASTTCRCPDFIPESHFLRVVNDSCFKILPMLDKSVDFIKAKASNGCV 240
DB 181 QONGIGVYVMASTTCRCPDFIPESHFLRVVNDSCFKILPMLDKSVDFIKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKERTTISPNNFLGQLIDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVYKERTTISPNNFLGQLIDYEKKIKN 300
QY 301 OTGASGPRKCLLHLKEKNEPVPVAVSBGQKSETPTLSPCADSATSEAGORPVHPASV 360
DB 301 OTGASGPRKCLLHLKEKNEPVPVAVSBGQKSETPTLSPCADSATSEAGORPVHPASV 360
QY 361 PSVPVQPSLLEDSPLVOALSGHLHSDRLSDSKLKRSSLDIKSVYSASVAAASLHGF 420
DB 361 PSVPVQPSLLEDSPLVOALSGHLHSDRLSDSKLKRSSLDIKSVYSASVAAASLHGF 420
QY 421 SSSREALEYRPTSTLLDGTNKLCOFSPVOELSEOTPEPTSPDKXEASIPKKLQTPARPSDQ 480
DB 421 SSSREALEYRPTSTLLDGTNKLCOFSPVOELSEOTPEPTSPDKXEASIPKKLQTPARPSDQ 480
QY 481 SKRLHVSATSSGTAQBSLSPLHRSQSVNDNHTSFLGLSTSOQHILTKSAGLGLKGMH 540
DB 481 SKRLHVSATSSGTAQBSLSPLHRSQSVNDNHTSFLGLSTSOQHILTKSAGLGLKGMH 540
QY 541 SDIAPOTSPSLTSSMYFATESSHFYASAIYGSASYSAYSCGOLPTCGOVYSVARR 600
DB 541 SDIAPOTSPSLTSSMYFATESSHFYASAIYGSASYSAYSCGOLPTCGOVYSVARR 600
QY 601 QKPSDRADSRHSWHESPFEKQPKRRSCOMERGEISIMENSRBELGKVGSSPFGSME 660
DB 601 QKPSDRADSRHSWHESPFEKQPKRRSCOMERGEISIMENSRBELGKVGSSPFGSME 660

DB 601 QKPSDRADSRHSWHESPFEKQPKRRSCOMERGEISIMENSRBELGKVGSSPFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665
RESULT 7
AAU75789
ID AAU75789 standard; protein; 665 AA.
AC AAU75789;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human protein phosphatase 7 (PP7) protein sequence.
XX
KW Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
KW acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
KW Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
KW dementia; Parkinson's disease; developmental disorder; Down's syndrome;
KW cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
KW melanoma; myeloma sarcoma.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT Domain 11..131
FT /label= Rhodanese_like_domain
FT 15..170
FT /note= "Dual specificity protein phosphatase"
FT 85..298
FT /note= "VH1-type dual specificity protein phosphatase"
FT 158..297
FT /label= Catalytic domain_DSPC
FT /note= "Dual specificity phosphatase"
FT 220..281
FT Region
FT /note= "Tyrosine specific protein phosphatases active
FT site tyr_phosphatase.prif"
FT 242..254
FT /note= "Tyrosine phosphatase"
FT 242..252
FT /note= "Tyrosine specific protein phosphatases"
PN WO200210363-A2.
XX
PD 07-FEB-2002.
XX
PF 26-JUL-2001; 2001WO-US023716.
XX
PR 28-JUL-2000; 2000US-0221679P.
PR 03-AUG-2000; 2000US-0223372P.
PR 10-AUG-2000; 2000US-0224309P.
PR 18-AUG-2000; 2000US-0226728P.
PR 30-AUG-2000; 2000US-0229254P.
PR 08-SEP-2000; 2000US-0231366P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Tang YT, Elliott VS, Raskumar J, Yao MG, Burford N, Wang YE;
PI Stewart BA, Gandhi AR, Patterson C, Lee BA, Hafalia AJA, Lu DAM;
PI Tribouley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;
PI Walla NK, Kearney L;
DR WPI; 2002-188735/24.
DR N-PSDB; ABX14474.
XX
PT New protein phosphatases, useful for diagnosing, treating or preventing
PT immune system disorders (e.g. Crohn's disease), neurological disorders
PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
PT cancers).
XX

PS Claim 1; Page 107-108; 117pp; English.

XX The present invention relates to a new polypeptide, a naturally occurring
CC amino acid sequence at least 95 % identical to it, a biologically active
CC fragment of it or an immunogenic fragment of it. The polypeptides,
CC polynucleotides, agonists and antagonists are useful for diagnosing,
CC treating or preventing disorders associated with aberrant expression of
CC protein phosphatases (PP), particularly immune system disorders e.g.
CC acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
CC asthma or Crohn's disease, neurological disorders e.g. epilepsy,
CC Huntington's disease, dementia or Parkinson's disease, developmental
CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
CC or sarcoma. The present amino acid sequence represents human protein
CC phosphatase 7 (PP7) which is one of several human protein phosphatases
CC (AAU5783-AAU5792) of the invention

XX Sequence 665 AA;

Query Match 99.6%; Score 3406; DB 5; Length 665;

Best Local Similarity 99.7%; Pred. No. 2.7e-256;

Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
DB 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
QY 61 DKVLITELIÖHSAKHKVIDCQKVVVYDQSSQDVASLSDCFLTVLGLKLEKSPNSVHL 120
DB 61 DKVLITELIÖHSAKHKVIDCQKVVVYDQSSQDVASLSDCFLTVLGLKLEKSPNSVHL 120
QY 121 LAGGFAFSPRCFPGLCGKSTLVPTCISQPCLPVANIPTRIIPNLYLGGQDVNLKELI 180
DB 121 LAGGFAFSPRCFPGLCGKSTLVPTCISQPCLPVANIPTRIIPNLYLGGQDVNLKELI 180
QY 181 QONGIGYVLNASTCPKPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKKAANGCV 240
DB 181 QONGIGYVLNASTCPKPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKKAANGCV 240
QY 241 LVHCLAGISRATIAIAYIMKRMDSLDEAYRFVKEKRTISPENFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRATIAIAYIMKRMDSLDEAYRFVKEKRTISPENFNLGQLLDYEKKIKN 300
QY 301 QTASGPKSKLKLHLKEKNEPVPAVSEGGQSETPSPCADSATSEAGORPVHPASV 360
DB 301 QTASGPKSKLKLHLKEKNEPVPAVSEGGQSETPSPCADSATSEAGORPVHPASV 360
QY 361 PSVPSPVQPSLLEDSPLVQALSGHLASDRLEDSNKLKRSFLDIKSVSYASMAASLHGF 420
DB 361 PSVPSPVQPSLLEDSPLVQALSGHLASDRLEDSNKLKRSFLDIKSVSYASMAASLHGF 420
QY 421 SSSSEDALEYKPSSTLLDGTNKLQFSPVQELSEQTETSPDKKEASIPKKLQIARPSDQ 480
DB 421 SSSSEDALEYKPSSTLLDGTNKLQFSPVQELSEQTETSPDKKEASIPKKLQIARPSDQ 480
QY 481 SKRLHSVRSSSGTAQRSLSLPLHRSGSVEDNHTSFLFGLSTSOQHLTKSAGLGLKGM 540
DB 481 SKRLHSVRSSSGTAQRSLSLPLHRSGSVEDNHTSFLFGLSTSOQHLTKSAGLGLKGM 540
QY 541 SDILAOTSTPSLTSSWYFATESSHFYSAIYIGSASVAYSCSOLPTCGDQVYVRRR 600
DB 541 SDILAOTSTPSLTSSWYFATESSHFYSAIYIGSASVAYSCSOLPTCGDQVYVRRR 600
QY 601 QKPSRADSRSSWHESSPEKQFKRRSCQMEFESIMSENRREELGKVQSSGFSGME 660
DB 601 QKPSRADSRSSWHESSPEKQFKRRSCQMEFESIMSENRREELGKVQSSGFSGME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 8

ABB97946
ID ABB97946 standard; protein; 665 AA.

XX ABB97946;

XX 06-SEP-2002 (first entry)

XX Human protein sequence #13.

XX Human; brain; tonsil; hippocampus; foetal brain; diagnosis.

XX Homo sapiens.

XX WO200252005-A1.

XX 04-JUL-2002.

XX 20-DEC-2001; 2001WO-JP011217.

XX 22-DEC-2000; 2000JP-00389742.

XX (KAZU-) KAZUSA DNA RES INST FOUND.

XX (CELE-) CELESTAR LEXICO-SCI LTD.

XX Ohara O, Nagase T, Nakajima D;

XX WPI, 2002-500762/53.

XX N-PSDB; ABB93966.

PT Genes and their expression products cloned from human cDNA libraries for
treatment and diagnosis of diseases associated with their expression.

PS Claim 1(a); Page 112-116; 238pp; Japanese.

CC The invention relates to DNA encoding polypeptides directly cloned from
CC cDNA libraries originating in adult whole brain, human tonsil, human
CC adult hippocampus and human foetal whole brain. Polypeptides and
CC polynucleotides of the invention may be used in the investigation of
CC differential expression of the DNA sequences in normal subjects and
CC disease patients. They may also be used in the production of antibodies,
CC oligonucleotide probes and DNA chips for diagnosis and identification of
CC drugs for treatment of diseases with which the DNA sequences are
CC associated. The sequences given in records ABB97934-A-BB97964 represent
CC human proteins of the invention

XX Sequence 665 AA;

Query Match 99.6%; Score 3406; DB 5; Length 665;

Best Local Similarity 99.7%; Pred. No. 2.7e-256;

Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
DB 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
QY 61 DKVLITELIÖHSAKHKVIDCQKVVVYDQSSQDVASLSDCFLTVLGLKLEKSPNSVHL 120
DB 61 DKVLITELIÖHSAKHKVIDCQKVVVYDQSSQDVASLSDCFLTVLGLKLEKSPNSVHL 120
QY 121 LAGGFAFSPRCFPGLCGKSTLVPTCISQPCLPVANIPTRIIPNLYLGGQDVNLKELI 180
DB 121 LAGGFAFSPRCFPGLCGKSTLVPTCISQPCLPVANIPTRIIPNLYLGGQDVNLKELI 180
QY 181 QONGIGYVLNASTCPKPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKKAANGCV 240
DB 181 QONGIGYVLNASTCPKPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKKAANGCV 240
QY 241 LVHCLAGISRATIAIAYIMKRMDSLDEAYRFVKEKRTISPENFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRATIAIAYIMKRMDSLDEAYRFVKEKRTISPENFNLGQLLDYEKKIKN 300
QY 301 QTASGPKSKLKLHLKEKNEPVPAVSEGGQSETPSPCADSATSEAGORPVHPASV 360
DB 301 QTASGPKSKLKLHLKEKNEPVPAVSEGGQSETPSPCADSATSEAGORPVHPASV 360

```

Db      301 QTGASGPKSKLHLLEKNEPVPVAVSEGGOKSTPLSPCCADSAITSEAGQRPVHPASV 360
Qy      361 PSVPVOPSLIEDSPVLOALSGHLMSADRLSDNSNKLKRSPLDIKSVYSASMAASLHG 420
Db      361 PSVPVOPSLIEDSPVLOALSGHLMSADRLSDNSNKLKRSPLDIKSVYSASMAASLHG 420
Qy      421 SSSDALERYKPSFTLLDGTNKLCOFSPVOCLESTQETSPDKKEMASIPKQLQTARPSDQ 480
Db      421 SSSDALERYKPSFTLLDGTNKLCOFSPVOCLESTQETSPDKKEMASIPKQLQTARPSDQ 480
Qy      481 SKRLHSVRTSSGTAQORSLSLPLHRSGVEDNYHTSFLFGSLSTQOHLTKSAGLGLKGMH 540
Db      481 SKRLHSVRTSSGTAQORSLSLPLHRSGVEDNYHTSFLFGSLSTQOHLTKSAGLGLKGMH 540
Qy      541 SDILAPQTSPLTSSMYFATESSHFYSAIYGSASASANSQCLPTCGDYVYVRR 600
Db      541 SDILAPQTSPLTSSMYFATESSHFYSAIYGSASASANSQCLPTCGDYVYVRR 600
Qy      601 QKPSDADRRSRHESHPKOPKRRSCMEFESIMSENRREELGKVSQSSFSGSM 660
Db      601 QKPSDADRRSRHESHPKOPKRRSCMEFESIMSENRREELGKVSQSSFSGSM 660
Qy      661 IIEVS 665
Db      661 IIEVS 665

```

```

RESULT 9
AAU79929
ID      AAU79929 standard; protein; 665 AA.
XX
XX
XX

```

```

AC      AAU79929;
XX
XX      02-JUL-2002 (first entry)
DT
DE      Human dual specificity phosphatase 21117 protein.
XX
XX      Human; dual specificity phosphatase 21117; erythroid-related disorder;
KW      haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
KV      erythrocytosis; liver-related disorder; cancer.
XX
OS      Homo sapiens.
XX
FH      Key
FH      Domain
FT      /label= Rhodanese_1like_domain
FT      21..24
FT      /note= "Casein kinase II phosphorylation site"
FT      91..94
FT      /note= "Casein kinase II phosphorylation site"
FT      158..297
FT      /note= "Dual specificity phosphatase catalytic domain"
FT      214..217
FT      /note= "Casein kinase II phosphorylation site"
FT      242..254
FT      /note= "Tyrosine specific protein phosphatase active
FT      site"
FT      242..254
FT      /note= "C-X5-R motif"
FT      266..269
FT      /note= "Casein kinase II phosphorylation site"
FT      369..372
FT      /note= "Casein kinase II phosphorylation site"
FT      421..424
FT      /note= "Casein kinase II phosphorylation site"
FT      434..437
FT      /note= "Casein kinase II phosphorylation site"
FT      458..461
FT      /note= "Casein kinase II phosphorylation site"
FT      508..511
FT      /note= "Casein kinase II phosphorylation site"
FT      589..592
FT      /note= "Casein kinase II phosphorylation site"
FT      Region

```

```

FT      Region
FT      /note= "Casein kinase II phosphorylation site"
FT      612..615
FT      Region
FT      /note= "Casein kinase II phosphorylation site"
FT      617..620
FT      Region
FT      /note= "Casein kinase II phosphorylation site"
FT      642..645
FT      /note= "Casein kinase II phosphorylation site"
PN      US2002034807-A1.
XX
XX      21-MAR-2002.
XX
XX      23-MAR-2001; 2001US-00816494.
XX
XX      24-MAR-2000; 2000US-0191858P.
XX      (MEYE/) MEYERS R A.
XX
XX      Meyers RA;
XX
XX      WPI; 2002-351086/38.
XX      N-PSDB; ABK49402.
XX
XX      New nucleic acids, designated 38692 and 21117, encoding dual specificity
XX      phosphatases for treating cell proliferation and differentiation
XX      disorders including hematopoietic and erythroid-related disorders and
XX      cancers.
XX
XX      Claim 8; Fig 1; 76pp; English.
XX
XX      The present invention relates to new nucleic acids designated 38692 and
XX      21117 encoding dual specificity phosphatase family members. The nucleic
XX      acid, polypeptide encoded by it, and antibody specific for the
XX      polypeptide may be used to diagnose and treat haematopoietic-related
XX      disorders such as leukaemias and autoimmune diseases, erythroid-related
XX      disorders and cancers, particularly of the breast, colon, adipose, prostate and
XX      lung. The present amino acid sequence represents the human dual
XX      specificity phosphatase 21117 protein of the invention, as described
XX      above
XX
SQ      Sequence 665 AA;

```

```

Qy      1 MAHEMIGQVTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSGLMRRIQQ 60
Db      1 MAHEMIGQVTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSGLMRRIQQ 60
Qy      61 DKVLITELIOMSAKHKVDIDCSQKVVVYDOSSODVASLSDCFVLVLLGKLEKSFNSVHL 120
Db      61 DKVLITELIOMSAKHKVDIDCSQKVVVYDOSSODVASLSDCFVLVLLGKLEKSFNSVHL 120
Qy      121 LAGGFAEFSCFPGLOEGKSTLVPTCISQCLPVANIGPTRILPMLYGCQRDVLANKELI 180
Db      121 LAGGFAEFSCFPGLOEGKSTLVPTCISQCLPVANIGPTRILPMLYGCQRDVLANKELI 180
Qy      181 QONGIGYVLANASYTCPPDFIPESHFLRPVNDSCETILPMWLDKSVDFLEKASNGCV 240
Db      181 QONGIGYVLANASYTCPPDFIPESHFLRPVNDSCETILPMWLDKSVDFLEKASNGCV 240
Qy      241 LVHCLAGISRSNTIATIMKRMDSUDEAVRPYKERRPTSPNFNFGQLDYEKKITN 300
Db      241 LVHCLAGISRSNTIATIMKRMDSUDEAVRPYKERRPTSPNFNFGQLDYEKKITN 300
Qy      301 QTGASGPKSKLHLLEKNEPVPVAVSEGGOKSTPLSPCCADSAITSEAGQRPVHPASV 360
Db      301 QTGASGPKSKLHLLEKNEPVPVAVSEGGOKSTPLSPCCADSAITSEAGQRPVHPASV 360
Qy      361 PSVPVOPSLIEDSPVLOALSGHLMSADRLSDNSNKLKRSPLDIKSVYSASMAASLHG 420
Db      361 PSVPVOPSLIEDSPVLOALSGHLMSADRLSDNSNKLKRSPLDIKSVYSASMAASLHG 420

```

QY 421 SSEDALERYKSTTLDGNTKLCQFSPVQELSEOTPETSPDKKEASIPKQLQTPARPSDSQ 480
 DB 421 SSEDALERYKSTTLDGNTKLCQFSPVQELSEOTPETSPDKKEASIPKQLQTPARPSDSQ 480
 QY 481 SKRLHSVRTSSSGCTAORSLSPLHRSGVVDNYHTSFLFGLSTSOQHLTKSAGLGKGMH 540
 DB 481 SKRLHSVRTSSSGCTAORSLSPLHRSGVVDNYHTSFLFGLSTSOQHLTKSAGLGKGMH 540
 QY 541 SDILAPQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDVYSVRRR 600
 DB 541 SDILAPQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDVYSVRRR 600
 QY 601 QKPSDRADSRWSHESPEKQFKRRSCOMEFGESIMSENRREELGKVGSSFFSGSME 660
 DB 601 QKPSDRADSRWSHESPEKQFKRRSCOMEFGESIMSENRREELGKVGSSFFSGSME 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665

RESULT 10

ABB97291
 ID ABB97291 standard; protein; 665 AA.

AC ABB97291;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 559.

KM Human; anti-neuritic; vulnary; anti-inflammatory; immunomodulator;
 KM anti-infective; cerebroprotective; cytoprotective; rheumatic; gene therapy;
 KM neuroprotective; anti-parkinsonian; protein therapy; EST;
 KM expressed sequence tag.

OS Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001MO-US026015.

PR 11-SEP-2000; 2000US-00659671.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Maundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Dirmann RT;
 DR N-PSDB; ABN32477.
 DR WPI; 2002-292408/33.

PT An isolated polynucleotide for treating diseases associated with its

PT encoded polypeptide such as cancer and multiple sclerosis.

PS Example 2; SEQ ID NO 559; 509pp; English.

CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 CC Sequence 665 AA;

Query Match 99.6%; Score 3406; DB 5; Length 665;
 Best local similarity 99.7%; Pred. No. 2,7e-256;
 Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTOIVTERVALLSEGTAKVLLIDSRPFVYNTSHIEAININCSKIMKRLQ 60
 DB 1 MAHEMIGTOIVTERVALLSEGTAKVLLIDSRPFVYNTSHIEAININCSKIMKRLQ 60
 QY 61 DKVLITELLQHSKAKVNDICQKVVVYDQSSQDVASLSSDCLTVLLGLTERKSPSVHL 120
 DB 61 DKVLITELLQHSKAKVNDICQKVVVYDQSSQDVASLSSDCLTVLLGLTERKSPSVHL 120
 QY 121 LAGGFAEFRCFPGLCCEKSTLVPCTISQPCLPVANIIGTRILPNIYLGCGQVRLKELM 180
 DB 121 LAGGFAEFRCFPGLCCEKSTLVPCTISQPCLPVANIIGTRILPNIYLGCGQVRLKELM 180
 QY 181 QONGIGYVLAASVTCBPDPFIPESHELRVPVNDSPCEKILPMWIDKSVDFLEKAKANGCV 240
 DB 181 QONGIGYVLAASVTCBPDPFIPESHELRVPVNDSPCEKILPMWIDKSVDFLEKAKANGCV 240
 QY 241 LVHCLAGISRSATIAIYIMKMDMSLDRAVRFVKEKRPITISPNFPLQGLDYEKIKX 300
 DB 241 LVHCLAGISRSATIAIYIMKMDMSLDRAVRFVKEKRPITISPNFPLQGLDYEKIKX 300
 QY 301 QTGASGPKSKLKLHLKEXNEBPVAVSEGGQKSETPSPPCADSATSEAGQRPVPAVY 360
 DB 301 QTGASGPKSKLKLHLKEXNEBPVAVSEGGQKSETPSPPCADSATSEAGQRPVPAVY 360
 QY 361 PSVPSVQPSLLEDSPLVQALSGHLASADRLLEDNKLKRSFSLDIKSVYSASMAASLHGF 420
 DB 361 PSVPSVQPSLLEDSPLVQALSGHLASADRLLEDNKLKRSFSLDIKSVYSASMAASLHGF 420
 QY 421 SSEDALERYKSTTLDGNTKLCQFSPVQELSEOTPETSPDKKEASIPKQLQTPARPSDSQ 480
 DB 421 SSEDALERYKSTTLDGNTKLCQFSPVQELSEOTPETSPDKKEASIPKQLQTPARPSDSQ 480
 QY 481 SKRLHSVRTSSSGCTAORSLSPLHRSGVVDNYHTSFLFGLSTSOQHLTKSAGLGKGMH 540
 DB 481 SKRLHSVRTSSSGCTAORSLSPLHRSGVVDNYHTSFLFGLSTSOQHLTKSAGLGKGMH 540
 QY 541 SDILAPQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDVYSVRRR 600
 DB 541 SDILAPQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDVYSVRRR 600
 QY 601 QKPSDRADSRWSHESPEKQFKRRSCOMEFGESIMSENRREELGKVGSSFFSGSME 660
 DB 601 QKPSDRADSRWSHESPEKQFKRRSCOMEFGESIMSENRREELGKVGSSFFSGSME 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665

RESULT 11

ADAS4744
 ID ADAS4744 standard; protein; 665 AA.

AC ADAS4744;

DT 20-NOV-2003 (first entry)

DE Human protein, SEQ ID 2312.

KM Cytostatic; Anti-inflammatory; Osteoprotective; Neuroprotective; Noctropic;
 KM Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KM inflammatory disease; osteoporosis; neurological disease.

OS Homo sapiens.

PN BPI293569-A2.

PD 19-MAR-2003.

XX

QY	601 QKSDPADSRSMWHESPPEKKFRRSCOMEFGESIMSENRREELAKVGSQSFSQSGME
Dd	601 QKSDRADSRSMWHESPPEKKFRRSCOMEFGESIMSENRREELAKVGSQSFSQSGME
QY	661 IIEVS 665
Dd	661 IIEVS 665
 RESULT 12	
ID	AAU79161 standard; protein; 665 AA.
XX	AAU79161;
DT	02-JUL-2002 (first entry)
DE	Human dual-specificity phosphatase-3 (DSP-16) mutant protein #1.
KW	Human, dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW	mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW	cancer; graft-versus-host disease; allergy; metabolic disease;
KW	abnormal cell growth; abnormal cell proliferation; contact inhibition;
KW	cell cycle abnormality; anchorage independent cell growth; apoptosis;
XX	intercellular adhesion; DSP-16 modulator; mutant; mutein.

Query Match	99.4%	Score 3399	DB 6	Length 665
Beat Local Similarity	99.5%	Pred. No. 9.4e-256		
Matches	662	Conservative 1	Mismatches 2	Indels 0
				Gaps 0
QY	1	MAHEMIGTQVTEETVALLSESGTEKVLIDSRPFVYNTMSTLLEAININCSKLMRRDQ	60	
Db	1	MAHEMIGTQVTEETVALLSESGTEKVLIDSRPFVYNTMSTLLEAININCSKLMRRDQ	60	
QY	61	DKVLITELLIGHSAKHKVIDDSCQKVYVYDSSODVASLSDCFVLVGLKLEKSPNVHL	120	
Db	61	DKVLITELLIGHSAKHKVIDDSCQKVYVYDSSODVASLSDCFVLVGLKLEKSPNVHL	120	
QY	121	LAGFAEFKRCFPGLCCEGKSTLVPTCISOPCLPVANITPRTLLPNLYGCGORDVLANELI	180	
Db	121	LAGFAEFKRCFPGLCCEGKSTLVPTCISOPCLPVANITPRTLLPNLYGCGORDVLANELI	180	
QY	181	QONGICYVLANASTYCKPDPFIESHFLEVPVNDSCCEKILPWLDKSYDFLEKXASNGCV	240	
Db	181	QONGICYVLANASTYCKPDPFIESHFLEVPVNDSCCEKILPWLDKSYDFLEKXASNGCV	240	
QY	241	LVHCLAGISRSATTIAIAYIMKRMDSLDEAARFYKEKRPRTSPNPNFLGOLLDYEEKIKTN	300	
Db	241	LVHCLAGISRSATTIAIAYIMKRMDSLDEAARFYKEKRPRTSPNPNFLGOLLDYEEKIKTN	300	
QY	301	QTGASGPKSKTLHLLEKPNPVPVAVSBGQKSTLPSPCADSATSSEAAGCPVPYPAV	360	
Db	301	QTGASGPKSKTLHLLEKPNPVPVAVSBGQKSTLPSPCADSATSSEAAGCPVPYPAV	360	
QY	361	PSVSVOPSLLEDSPVVALSGHLISADRLIEDNSKTKRSSTLDIKSVSASMAASLHGF	420	
Db	361	PSVSVOPSLLEDSPVVALSGHLISADRLIEDNSKTKRSSTLDIKSVSASMAASLHGF	420	
QY	421	SSSDALEEYKPESTLIDGTNKLCOFSPVDELSTQETBSPDXBEASIPKLTQTAPPSNQ	480	
Db	421	SSSDALEEYKPESTLIDGTNKLCOFSPVDELSTQETBSPDXBEASIPKLTQTAPPSNQ	480	
QY	481	SKRLHSVRTSSSGTAQRSLSLPLHRSAGSYEDNYHNSFLPGSTSOOHLTKSAGLKGWH	540	
Db	481	SKRLHSVRTSSSGTAQRSLSLPLHRSAGSYEDNYHNSFLPGSTSOOHLTKSAGLKGWH	540	
QY	541	SDLIAPOTSPSLTSSWFATSSSHYTSASAIYGGASASAVSCGOLPTCGDQYVYVRR	600	
Db	541	SDLIAPOTSPSLTSSWFATSSSHYTSASAIYGGASASAVSCGOLPTCGDQYVYVRR	600	

QY	601	OKSDDADSRSSWHEEPPEKQFKRRSCOMEFGESIMSNRREELAKVGSOSGFSQSMTE
Dd	601	OKSDRADRSRWHEERSPPKEQFKRRSCOMFEGSINSNRRLKVKVSGSSFFSQSMTE
QY	661	IIEVS 665
Dd	661	IIEVS 665
RESULT 12		
AAU79161		
ID	AAU79161	strand; protein, 665 AA.
XX	AAU79161;	
AC		
XX		
DT	02-JUL-2002	(first entry)
DE		
XX		
KM	Human dual-specificity phosphatase-3 (DSP-16) mutant protein #1.	
KW	Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;	
KM	mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;	
KW	cancer; graft-versus-host disease; allergy; metabolic disease;	
KM	abnormal cell growth; abnormal cell proliferation; contact inhibition;	
KW	cell cycle abnormality; anchorage independent cell growth; apoptosis;	
KM	intercellular adhesion; DSP-16 modulator; mutant; mutein.	
OS		
OS	Homo sapiens.	
OS	Synthetic.	
XX		
PH	Key	Location/Qualifiers
FT	Misc-difference 213	/note= "Wild-type Asp substituted by Ala"
FT		
XX	W0200226997-A2.	
PN		
XX	04-APR-2002.	
PD		
PX	25-SEP-2001; 2001WO-US030124.	
PR		
XX	26-SEP-2000; 2000US-0235487P.	
PA	(CEPT-) CEPTYR INC.	
PI		
Luche RM, Wei B;		
DR	WPI; 2002-315802/35.	
PT	New DSP-16 polypeptide, useful for identifying modulators of its	
PT	activity, which can be used in the treatment of disorders such as	
XX	Duchenne muscular dystrophy, or cancer.	
PS	Claim 46; Page; 87pp; English.	
XX		
XX	The present invention relates to a new polypeptide, DSP-16, having a 665	
CC	amino acid sequence, given in the specification, or a variant having at	
CC	least 50 % identical residues, which retains the ability to	
CC	dephosphorylate an activated mitogen-activated protein (MAP) kinase. The	
CC	invention can be used for identifying agents which modulate DSP-16	
CC	activity, for modulation of a proliferative response in a cell, survival	
CC	of a cell, or differentiation of a cell. The cell displays contact	
CC	inhibition of cell growth or anchorage independent growth and may display	
CC	altered intercellular adhesion. The agent may modulate apoptosis, or the	
CC	cell cycle. The identified modulators can be used to treat Duchenne	
CC	muscular dystrophy, cancer, graft-versus-host disease, autoimmune	
CC	diseases, allergies, metabolic diseases, abnormal cell growth, abnormal	
CC	cell proliferation, and cell cycle abnormalities. The present amino acid	
CC	sequence represents the human dual-specificity phosphatase-3 (DSP-16)	
CC	mutant protein #1. Note: This sequence is not shown in the specification	
CC	but is derived from the wild-type human DSP-16 (AAU79156) protein given	
CC	in figure 2 of the specification	
XX		
Sequence 665 AA;		
QY		

Query Match 99.4%; Score 3398; DB 5; Length 665;
 Best Local Similarity 99.5%; Pred. No. 1,1e-255;
 Matches 662; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMIGTOIVTERLVALLESGETEKVLLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 DB 1 MAHEMIGTOIVTERLVALLESGETEKVLLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 QY 61 DKVLITELIQHSKHKVDIDCSQKVVYDSSQDVASLSDCFLTVLLGLKLEKSFNSVHL 120
 DB 61 DKVLITELIQHSKHKVDIDCSQKVVYDSSQDVASLSDCFLTVLLGLKLEKSFNSVHL 120
 QY 121 LAGGFAEFSRCPGLCEGKSTLVPTCISOPCLPVANIGPRTIIPNLYLGCQDVLNKEIM 180
 DB 121 LAGGFAEFSRCPGLCEGKSTLVPTCISOPCLPVANIGPRTIIPNLYLGCQDVLNKEIM 180
 QY 181 QONGIGVYLNASTCPKPDPIPSHFLRPVNDSPCEKILPMDKSVDFLEKAKASNGCV 240
 DB 181 QONGIGVYLNASTCPKPDPIPSHFLRPVNDSPCEKILPMDKSVDFLEKAKASNGCV 240
 QY 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKN 300
 DB 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKN 300
 QY 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGQKSETPLSPCADSATSEAGORPVHPASV 360
 DB 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGQKSETPLSPCADSATSEAGORPVHPASV 360
 QY 361 PSVPVQPSLLEDSPLVQALSGHLISADRLSDNKLKRSFSLDIKSVYSASMAASLHGF 420
 DB 361 PSVPVQPSLLEDSPLVQALSGHLISADRLSDNKLKRSFSLDIKSVYSASMAASLHGF 420
 QY 421 SSSBDALEYKPPSTTLDGINKLCPSPVQELSEQPTPTSPDKBASIPKQLQTPARPSDQ 480
 DB 421 SSSBDALEYKPPSTTLDGINKLCPSPVQELSEQPTPTSPDKBASIPKQLQTPARPSDQ 480
 QY 481 SKRLHSVRTSSSGAQRSLSPHRSQSVEDNYHTSLFGLSTSQOQLTSAGLKGWH 540
 DB 481 SKRLHSVRTSSSGAQRSLSPHRSQSVEDNYHTSLFGLSTSQOQLTSAGLKGWH 540
 QY 541 SDLIAPQTSPTSLTSSWYFATSSSHFYSASAIYGSASAYAGSQCPLTGDDQVYVRRR 600
 DB 541 SDLIAPQTSPTSLTSSWYFATSSSHFYSASAIYGSASAYAGSQCPLTGDDQVYVRRR 600
 QY 601 QKPSDRAISRWSHESPFEPKQFRRSCQMEFGESINENRSBELGKVSQSPSGSME 660
 DB 601 QKPSDRAISRWSHESPFEPKQFRRSCQMEFGESINENRSBELGKVSQSPSGSME 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665

RESULT 13
 AAU79162 standard; protein; 665 AA.
 ID AAU79162;
 AC AAU79162;
 XX 02-JUL-2002 (first entry)
 DE Human dual-specificity phosphatase-3 (DSP-16) mutant protein #2.
 KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KW cancer; graft-versus-host disease; allergy; metabolic disease;
 KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KW intercellular adhesion; DSP-16 modulator; mutant; mutein.
 KM Homo sapiens.
 OS Synthetic.
 XX

Key Location/Qualifiers
 FT Misc-difference 244 /note= "Wild-type Cys substituted by Ser"
 PN MO200226997-A2.
 PD 04-APR-2002.
 PF 25-SEP-2001, 2001MO-US030124.
 PR 26-SEP-2000, 2000US-0235487P.
 PA (CEPT-) CEPTYR INC.
 PI Luche RM, Wei B;
 PS WPI; 2002-315802/35.
 DR New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 XX Claim 46; Page; 87pp; English.
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity, for modulation of a proliferative response in a cell, survival
 CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present amino acid
 CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 CC mutant protein #2. Note: This sequence is not shown in the specification
 CC but is derived from the wild-type human DSP-16 (AAU79156) protein given
 CC in figure 2 of the specification
 XX
 SQ Sequence 665 AA;

Query Match 99.4%; Score 3396; DB 5; Length 665;
 Best Local Similarity 99.5%; Pred. No. 1,1e-255;
 Matches 662; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMIGTOIVTERLVALLESGETEKVLLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 DB 1 MAHEMIGTOIVTERLVALLESGETEKVLLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 QY 61 DKVLITELIQHSKHKVDIDCSQKVVYDSSQDVASLSDCFLTVLLGLKLEKSFNSVHL 120
 DB 61 DKVLITELIQHSKHKVDIDCSQKVVYDSSQDVASLSDCFLTVLLGLKLEKSFNSVHL 120
 QY 121 LAGGFAEFSRCPGLCEGKSTLVPTCISOPCLPVANIGPRTIIPNLYLGCQDVLNKEIM 180
 DB 121 LAGGFAEFSRCPGLCEGKSTLVPTCISOPCLPVANIGPRTIIPNLYLGCQDVLNKEIM 180
 QY 181 QONGIGVYLNASTCPKPDPIPSHFLRPVNDSPCEKILPMDKSVDFLEKAKASNGCV 240
 DB 181 QONGIGVYLNASTCPKPDPIPSHFLRPVNDSPCEKILPMDKSVDFLEKAKASNGCV 240
 QY 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKN 300
 DB 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKN 300
 QY 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGQKSETPLSPCADSATSEAGORPVHPASV 360
 DB 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGQKSETPLSPCADSATSEAGORPVHPASV 360
 QY 361 PSVPVQPSLLEDSPLVQALSGHLISADRLSDNKLKRSFSLDIKSVYSASMAASLHGF 420

```

Db      361 PSVPSVOPSLIEDSPVQALSGILHLSADRLIEDSNKLSFSLDIKSVSYSAWMAASLHGF 420
QY      421 SSEDALLEYKPTTLDGNTNKLCOFSPVOELSEQTPETSPPDEKASIPKKTOTARPDSQ 480
Db      421 SSEDALLEYKPTTLDGNTNKLCOFSPVOELSEQTPETSPPDEKASIPKKTOTARPDSQ 480
QY      481 SKRLHSVRTSSSGTQARSLISPLHRSGSVEDNYHTSFLGLSTSOQHLTKSAGLGLKGMH 540
Db      481 SKRLHSVRTSSSGTQARSLISPLHRSGSVEDNYHTSFLGLSTSOQHLTKSAGLGLKGMH 540
QY      541 SDILAPQSTPSTLTSWYFATESSHFYASATYGGASASVASCQLPCCGQVYSVRR 600
Db      541 SDILAPQSTPSTLTSWYFATESSHFYASATYGGASASVASCQLPCCGQVYSVRR 600
QY      601 QKPSDRADSRSRWHESSPPEKQFKRRSCOMERGESIMSENREELGKVGSSPFGSME 660
Db      601 QKPSDRADSRSRWHESSPPEKQFKRRSCOMERGESIMSENREELGKVGSSPFGSME 660
QY      661 IIEVS 665
Db      661 IIEVS 665

```

RESULT 14

ABR52352 standard; protein; 665 AA.

ABR52352;

19-JUN-2003 (first entry)

Protein relating to the invention SEQ ID NO: 42.

antiapoptotic; hepatotropic; nephrotropic; antiarthritic;
 antiproliferative; cardiatic; cytoskeletal; gene therapy; liver disease;
 proliferative disorder; renal failure; cardiovascular disorder;
 immunological disorder; arthritis; psoriasis; congenital heart defect;
 congestive heart failure; cancer; human; mouse; rat; fruitfly.

Homo sapiens.

WO200257460-A2.

25-JUN-2002.

20-DEC-2001; 2001WO-US050459.

20-DEC-2001; 2000US-0256868P.

30-MAR-2001; 2001US-0280186P.

01-MAY-2001; 2001US-0287735P.

05-JUN-2001; 2001US-0295848P.

25-JUN-2001; 2001US-0300465P.

(BRIM) BRISTOL-MYERS SQUIBB CO.

JACKSON DG, Feder J, Nelson T, Mintler G, Ramanathan C, Lee L,

Siemens N, Bol D, Schieven G, Finger J, Rodderud CG, Bassolino D,

Krystek S, Mcatee P, Suchard S, Banae D;

WPI; 2002-599721/64.

N-PSDB; ACC60521.

Novel polynucleotides encoding human phosphatase polypeptides useful in

the prevention or treatment of e.g. proliferative and cardiovascular

disorders.

Claim 5; Fig 5; 801pp; English.

The invention relates to a novel isolated nucleic acid comprising a

polynucleotide having a nucleotide sequence selected from 40

polynucleotides fully defined in the specification. The polynucleotide of

the invention has antiapoptotic, hepatotropic, nephrotropic,

antiarthritic, antiproliferative, cardiatic, and cytoskeletal activity. The
 polynucleotide may have a use in gene therapy. A polynucleotide or
 amelioreting a medical condition, e.g. a proliferative, creating or
 also useful for treating e.g. liver disease, renal failure, immunological
 disorders including arthritis and psoriasis, cardiovascular disorders
 such as congenital heart defects and congestive heart failure, and
 cancer. A method of the invention is useful for diagnosing a pathological
 condition or susceptibility to a condition in a subject. The present
 sequence is used in the exemplification of the invention

Sequence 665 AA;

Query Match 99.1%; Score 3380; DB 5; Length 665;

Best Local Similarity 99.4%; Pred. No. 6, 7e-255; Matches 661; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 MAHEMIGTQIVTERLVALLESSTGTEKYLIDSRPFVEYNTSHILEAININCKLMKRLQ 60
 1 MAHEMIGTQIVTERLVALLESSTGTEKYLIDSRPFVEYNTSHILEAININCKLMKRLQ 60

61 DKVILTELIOHSAKHVDIDSCQVYVVOSSQDVASISDCLTYLGLKESFNSVHL 120
 61 DKVILTELIOHSAKHVDIDSCQVYVVOSSQDVASISDCLTYLGLKESFNSVHL 120

121 LAGGPAEFRCFPGLECKSTLPTCISQPCLVANIGFTRILPNLYLGCQDVANKELI 180
 121 LAGGPAEFRCFPGLECKSTLPTCISQPCLVANIGFTRILPNLYLGCQDVANKELI 180

181 QONGIGVYANASTYCKPDPFIPSHFLRVVNDSCFKILPWIDSVDFEAKASNGCV 240
 181 QONGIGVYANASTYCKPDPFIPSHFLRVVNDSCFKILPWIDSVDFEAKASNGCV 240

241 LVHCLAGISRSATIAIYIMKMDSLDAVYFVEKERTISPNFNGOLLDEKTKN 300
 241 LVHCLAGISRSATIAIYIMKMDSLDAVYFVEKERTISPNFNGOLLDEKTKN 300

301 QTGASGPKSKLKLHLEKNEBPVAVSEGOQSETPPLSPCADSATSEMAQORPVHPASV 360
 301 QTGASGPKSKLKLHLEKNEBPVAVSEGOQSETPPLSPCADSATSEMAQORPVHPASV 360

361 PSVPSVOPSLIEDSPVQALSGILHLSADRLIEDSNKLSFSLDIKSVSYSAWMAASLHGF 420
 361 PSVPSVOPSLIEDSPVQALSGILHLSADRLIEDSNKLSFSLDIKSVSYSAWMAASLHGF 420

421 SSEDALLEYKPTTLDGNTNKLCOFSPVOELSEQTPETSPPDEKASIPKKTOTARPDSQ 480
 421 SSEDALLEYKPTTLDGNTNKLCOFSPVOELSEQTPETSPPDEKASIPKKTOTARPDSQ 480

481 SKRLHSVRTSSSGTQARSLISPLHRSGSVEDNYHTSFLGLSTSOQHLTKSAGLGLKGMH 540
 481 SKRLHSVRTSSSGTQARSLISPLHRSGSVEDNYHTSFLGLSTSOQHLTKSAGLGLKGMH 540

541 SDILAPQSTPSTLTSWYFATESSHFYASATYGGASASVASCQLPCCGQVYSVRR 600
 541 SDILAPQSTPSTLTSWYFATESSHFYASATYGGASASVASCQLPCCGQVYSVRR 600

601 QKPSDRADSRSRWHESSPPEKQFKRRSCOMERGESIMSENREELGKVGSSPFGSME 660
 601 QKPSDRADSRSRWHESSPPEKQFKRRSCOMERGESIMSENREELGKVGSSPFGSME 660

661 IIEVS 665
 661 IIEVS 665

RESULT 15
 ABR52424 standard; protein; 664 AA.

ABR52424;

19-JUN-2003 (first entry)

XX Protein relating to the invention SEQ ID NO: 190.

DE anti-proliferative; hepatotropic; nephrotropic; anti-arthritic;
XX anti-proliferative; cardiatic; cytostatic; gene therapy; liver disease;
KM anti-proliferative disorder; renal failure; cardiovascular disorder;
KM immunological disorder; arthritis; psoriasis; congenital heart defect;
KM congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.

OS Homo sapiens.

XX WO200257460-A2.

XX 25-JUL-2002.

XX 20-DEC-2001; 2001WO-US050459.

XX 20-DEC-2000; 2000US-0256868P.

XX 30-MAR-2001; 2001US-0280186P.

XX 01-MAY-2001; 2001US-0287735P.

XX 05-JUN-2001; 2001US-0295848P.

XX 25-JUN-2001; 2001US-0300465P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Jackson DQ, Feder J, Nelson T, Muntier G, Ramanathan C, Lee L;

XX Stiemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;

XX Klystek S, Mcatee P, Suchard S, Banas D;

XX WPI; 2002-599721/64.

XX Novel polynucleotides encoding human phosphatase polypeptides useful in

XX the prevention or treatment of e.g. proliferative and cardiovascular

XX disorders.

XX Example 57; Page 500-501; 801pp; English.

XX The invention relates to a novel isolated nucleic acid comprising a

XX polynucleotide having a nucleotide sequence selected from 40

XX polynucleotides fully defined in the specification. The polynucleotide of

XX the invention has anti-proliferative, hepatotropic, nephrotropic,

XX anti-arthritic, anti-psoriatic, cardiatic, and cytostatic activity. The

XX polynucleotide may have a use in gene therapy. A polynucleotide or

XX polypeptide of the invention is useful for preventing, treating or

XX ameliorating a medical condition, e.g. a proliferative disorder. They are

XX also useful for treating e.g. liver disease, renal failure, immunological

XX disorders including arthritis and psoriasis, cardiovascular disorders

XX such as congenital heart defects and congestive heart failure, and

XX cancer. A method of the invention is useful for diagnosing a pathological

XX condition or susceptibility to a condition in a subject. The present

XX sequence is used in the exemplification of the invention

XX Sequence 664 AA;

Query Match 98.6%; Score 3368.5; DB 5; Length 664;

Best Local Similarity 99.1%; Pred. No. 2.2e-253;

Matches 659; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 MAHEMIGTQVTEVTLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
DB 1 MAHE-IGTQVTEVTLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 59
QY 61 DKVLITELIQHSAGKXVDICDQKVVYDDSSQDVASLSDCFLTVLLGKLEKSFNSVHL 120
DB 60 DKVLITELIQHSAGKXVDICDQKVVYDDSSQDVASLSDCFLTVLLGKLEKSFNSVHL 119
QY 121 LAGGFAEFSRCPGLGCKSTLVPTCTSQPCLPVANIGPRTIIPNLVYGCQDVLANKELEI 180
DB 120 LAGGFAEFSRCPGLGCKSTLVPTCTSQPCLPVANIGPRTIIPNLVYGCQDVLANKELEI 179
QY 181 QONGIGVYLANASYCPKPDFIPESHFLRPVNDSECEKILPMLDKSVDFIEKAKANGCV 240
DB 180 QONGIGVYLANASYCPKPDFIPESHFLRPVNDSECEKILPMLDKSVDFIEKAKANGCV 239

QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRPFVKEKRPITISPNFPLGOLLPEKKIKN 300
DB 240 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRPFVKEKRPITISPNFPLGOLLPEKKIKN 299
QY 301 QTSAGSPKSKLKLHLEKNEPVPVAVSEGQKSETPSPCADSATSEAAQRPVHPASV 360
DB 300 QNASGPKSKLKLHLEKNEPVPVAVSEGQKSETPSPCADSATSEAAQRPVHPASV 359
QY 361 PSVPSVQPSLLEDSPVLVQALSGHLSDADRLSDSNKLRSPSLDIKSVYSASMAASLHGF 420
DB 360 PSVPSVQPSLLEDSPVLVQALSGHLSDADRLSDSNKLRSPSLDIKSVYSASMAASLHGF 419
QY 421 SSEDALFYKPPETTLTGNTKLCQPSPVQELSPQTPETSFDKEASIPKLLQTPAPSDQ 480
DB 420 SSEDALFYKPPETTLTGNTKLCQPSPVQELSPQTPETSFDKEASIPKLLQTPAPSDQ 479
QY 481 SKRLHSVRTSSSGTAQRSLSPHRSQSVEDNYHTSFLGLSTSOQHLTKSAGLGKGMH 540
DB 480 SKRLHSVRTSSSGTAQRSLSPHRSQSVEDNYHTSFLGLSTSOQHLTKSAGLGKGMH 539
QY 541 SDILAPQSTPSTLSSWYFATESSHFYASASAIYGSASYSAYSCQLPTCGDOVYSVRRR 600
DB 540 SDILAPQSTPSTLSSWYFATESSHFYASASAIYGSASYSAYSCQLPTCGDOVYSVRRR 599
QY 601 QKPSDADSRKSWHESPPKQPKRRSCOMEFESIMSENRSPREELGXVGSOSFSGSME 660
DB 600 QKPSDADSRKSWHESPPKQPKRRSCOMEFESIMSENRSPREELGXVGSOSFSGSME 659
QY 661 IIEVS 665
DB 660 IIEVS 664

Search completed: June 21, 2004, 13:20:33
Job time : 76.2709 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2004, 14:15:15 ; Search time 1 Seconds
(without alignments)
1.329 Million cell updates/sec

Title: US-10-029-345A-109
Perfect score: 3418
Sequence: 1 MAHEMIGTOIVTERLVALLE.....LGRVGSQSSFGSGMERIEVS 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3 seqs, 1998 residues

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 3 summaries

Database : rag109.pdp.*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3057.5	89.5	660	1 ABR52385	Protein relating t
2	3030.5	88.7	666	1 AAB20325	Human protein phos
3	2930	85.7	672	1 AAM25744	Human protein sequ

ALIGNMENTS

RESULT 1
ID ABR52385
ABR52385 standard; protein; 660 AA.
AC ABR52385;
XX
XX
DT 19-JUN-2003 (first entry)
XX
XX
DE Protein relating to the invention SEQ ID NO: 114.
XX
XX
XX anti-proliferative; hepatotropic; nephrotropic; anti-arthritic;
XX anti-ischaemic; cardiac; cytosolic; gene therapy; liver disease;
XX proliferative disorder; renal failure; cardiovascular disorder;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
XX Mus musculus.
XX
XX PN MO200257460-A2.
XX
XX PD 25-JUL-2002.
XX
XX PF 20-DEC-2001; 2001MO-US050459.
XX
XX PR 20-DEC-2000; 2000US-0256868P.

PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Jackson DG, Feder J, Nelson T, Minter G, Ramanathan C, Lee L,
PI Siemsen N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
PI Kystek S, Mcatee P, Suchard S, Banas D;

XX WPI; 2002-599721/64.
DR N-PSDB; ACC60560.

PT Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.

XX Disclosure; Fig 12; 801pp; English.

CC The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has anti-proliferative, hepatotropic, nephrotropic,
CC anti-arthritic, anti-ischaemic, cardiac, and cytosolic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention

XX SQ Sequence 660 AA;

Query Match 89.5%; Score 3057.5; DB 1; Length 660;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 594; Conservative 29; Mismatches 37; Indels 5; Gaps 3;

QY 1 MAHEMIGTOIVTERLVALLESGTEKYLIDSRPFVYNTSHILEAININCKMKRRLOQ 60
DB 1 MAHEMIGTOIVTERLVALLESGTEKYLIDSRPFVYNTSHILEAININCKMKRRLOQ 60
QY 61 DKVLTITELHQSHAKKAVDIDCSQKVVVYDDSSQDVASLSDCFLTYLGLKESFNSVHL 120
DB 61 DKVLTITELHQSHAKKAVDIDCSQKVVVYDDSSQDVASLSDCFLTYLGLKESFNSVHL 120
QY 121 LAGGFAEFNRCPGLCEGKSTLVPTCISOPCLPVANIGPRIIPNLYLGGQRPVLANELI 180
DB 121 LAGGFAEFNRCPGLCEGKSTLVPTCISOPCLPVANIGPRIIPNLYLGGQRPVLANELI 180
QY 181 QQNGIGVYNASVTCRKPDPFIPESHPLRVVNDSPCEKILPWLKSVDFTEKASNGCV 240
DB 181 QQNGIGVYNASVTCRKPDPFIPESHPLRVVNDSPCEKILPWLKSVDFTEKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRVVKKRPTISNPNFLGGLDLYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRVVKKRPTISNPNFLGGLDLYEKKIKN 300
QY 301 QTSASGPKSKLKLHLKENEPEVPAVSEGGQKSETPISPPCADSATSEAGRPVHPASV 360
DB 301 QTSASGPKSKLKLHLKENEPEVPAVSEGGQKSETPISPPCADSATSEAGRPVHPASV 360
QY 361 PSYPSVQPSLLEBDSPLVQALSGHLASDRLSDSNKLKRSFSLDKSVSYASMAASLHGF 420
DB 361 PSYPSVQPSLLEBDSPLVQALSGHLASDRLSDSNKLKRSFSLDKSVSYASMAASLHGF 420
QY 421 SSGEDALEYKPSBTLDGNTKLQPSVQVQELSGTPESTPDKEAATPKLQTPARSDSQ 480
DB 421 SSGEDALEYKPSBTLDGNTKLQPSVQVQELSGTPESTPDKEAATPKLQTPARSDSQ 480
QY 475 -SSEALDYCKPSATLDGNTKLQPSVQVQVSESSPETSPEKEAHI PKQPPRPSBSQ 475
DB 475 -SSEALDYCKPSATLDGNTKLQPSVQVQVSESSPETSPEKEAHI PKQPPRPSBSQ 475

QY	481	SVKRLHSVFRSSSGTARSLILSPHREGSVEDNDHTNFTFLPGLSTNSQOHLTYSAGLGLKGMH	540
Db	476	VTRLHSVRKSGSSOSTKRPFFSPHLRSGSVEDNDHTNFTFLPGLSTNSQOHLTYSAGLGLKGMH	535
QY	541	SDILAPQSTPSLTSSWYFATBESSHFSYASAIYGSASYSAYSCSLPTCGDQVYSVRR	600
Db	536	SDILAPQSGAPSLTSSWYFATBESHLYASAIYGNSSYSAYSCGQLPTCSDDIYVRRR	595
QY	601	QKPSPRADRRSRWHSESPPEKQPKRRSCOMEFEESITMSNRSREELGRVSGSGSFGSGME	660
Db	556	QKPTRADRRSRWHSESPPEKQPKRRSCOMEFEESITMSNRSREELGRVSGSGSFGSGME	655
QY	661	IIEVSS	665
Db	656	IIEVSS	660
RESULT 2			
ID	AAB20325		
XX	AAB20325 standard; protein; 666 AA.		
XX	AAB20325;		
DT	29-MAY-2001 (first entry)		
DE	Human protein phosphatase and kinase protein-4.		
XX			
KW	Protein phosphatase and kinase protein; PPKR-4; human;		
KM	gastrointestinal disorder; immune system disorder; neurological disorder;		
XX	cell proliferative disorder; cancer; diagnosis; therapy.		
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	Modified-site	12	/note= "O-phosphorylated"
FT	Modified-site	21	/note= "O-phosphorylated"
FT	Modified-site	23	/note= "O-phosphorylated"
FT	Modified-site	38	/note= "O-phosphorylated"
FT	Modified-site	38	/note= "N-glycosylated"
FT	Modified-site	49	/note= "N-glycosylated"
FT	Modified-site	72	/note= "N-glycosylated"
FT	Modified-site	82	/note= "O-phosphorylated"
FT	Region	85..298	/note= "O-phosphorylated"
FT	Modified-site	91	/note= "VH1-type dual specificity phosphatase signature"
FT	Modified-site	190	/note= "O-phosphorylated"
FT	Modified-site	212	/note= "N-glycosylated"
FT	Modified-site	214	/note= "N-glycosylated"
FT	Active-site	220..280	/note= "O-phosphorylated"
FT	Region	237..278	/note= "tyrosine specific protein phosphatase"
FT	Modified-site	266	/note= "Y phosphatase signature"
FT	Modified-site	280	/note= "O-phosphorylated"
FT	Modified-site	300	/note= "O-phosphorylated"
FT	Modified-site	369	/note= "N-glycosylated"
FT	Modified-site	393	/note= "O-phosphorylated"

FT	Modified-site	/note= "O-phosphorylated"
FT	421	
FT	Modified-site	/note= "O-phosphorylated"
FT	422	
FT	Modified-site	/note= "O-phosphorylated"
FT	434	
FT	Modified-site	/note= "O-phosphorylated"
FT	439	
FT	Modified-site	/note= "O-phosphorylated"
FT	468	
FT	Modified-site	/note= "O-phosphorylated"
FT	471	
FT	Modified-site	/note= "O-phosphorylated"
FT	479	
FT	Modified-site	/note= "O-phosphorylated"
FT	528	
FT	Modified-site	/note= "O-phosphorylated"
FT	590	
FT	Modified-site	/note= "O-phosphorylated"
FT	597	
FT	Modified-site	/note= "O-phosphorylated"
FT	605	
FT	Modified-site	/note= "O-phosphorylated"
FT	610	
FT	Modified-site	/note= "O-phosphorylated"
FT	613	
FT	Modified-site	/note= "O-phosphorylated"
FT	618	
FT	Modified-site	/note= "O-phosphorylated"
FT	628	
FT	Modified-site	/note= "O-phosphorylated"
FT	641	
FT	Modified-site	/note= "N-glycosylated"
FT	643	
FT	Modified-site	/note= "O-phosphorylated"
FN	W0200120004-A2.	
XX		
XX	22-MAR-2001.	
XX		
XX	14-SEP-2000; 2000MO-US025515.	
XX		
XX	15-SEP-1999; 99US-0154141P.	
PR		
XX	(INCY-) INCYTE GENOMICS INC.	
PA		
XX		
XX	Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y;	
PI	Lu DM;	
XX		
XX	WPI; 2001-244811/25.	
DR	N-PSDB; AAF30479.	
XX		
PT	Novel human protein phosphatase and kinase proteins for diagnosis,	
PT	treatment and prevention of gastrointestinal, immune system, neurological	
PT	and cell proliferative disorders.	
XX		
XX	Claim 1; Page 87-88; 103pp; English.	
PS		
XX		
XX	The present sequence is that of novel human protein phosphatase and	
CC	kinase protein PPHK-4, as predicted from Incyte Clone ID No. 1234795CB1	
CC	(see AAF30479). Tissues that express PPHK-4 (as a fraction of total	
CC	tissues expressing PPHK-4) include gastrointestinal (0.385),	
CC	cardiovascular (0.154), nervous (0.154) and reproductive (0.154).	
CC	Diseases or conditions associated with tissues expressing PPHK-4 (as a	
CC	fraction of total tissues expressing PPHK-4) include cancer (0.692),	
CC	inflammation or trauma (0.308) and cell proliferation (0.231). The	
CC	encoded protein shows homology to mouse neuronal tyrosine threonine	
CC	phosphatase 1. The invention provides human PPHK-1 to -11 polypeptides	
CC	(see AAF30476-86). It also provides	
CC	expression vectors, host cells, antibodies, agonists and antagonists, as	
CC	well as methods for diagnosing, treating or preventing disorders	
CC	associated with expression of PPHK, including gastrointestinal	
CC	disorders, immune system disorders, neurological disorders and cell	

CC proliferative disorders, including cancer
XX Sequence 666 AA;

Query Match 88.7%; Score 3030.5; DB 1; Length 666;
Best Local Similarity 88.9%; Pred. No. 0;
Matches 607; Conservative 8; Mismatches 33; Indels 35; Gaps 5;

QY 1 MAHEMIGTOIVTERLVALESGETEKVLLIDSRPFVEVNTSHLEAININCKMKRRLQ 60
DB 1 MAHEMIGTOIVTERLVALESGETEKVLLIDSRPFVEVNTSHLEAININCKMKRRLQ 60
QY 61 DKVLITELIQSHAKHVIDIDCQKVVVYDQSSQDVASLSDDCEFLVYLGLKLEKSPNSVHL 120
DB 61 DKVLITELIQSHAKHVIDIDCQKVVVYDQSSQDVASLSDDCEFLVYLGLKLEKSPNSVHL 120
QY 121 LAGGFAPFRCFPGLECKGSTLVPTCISOPCLPVANIGPRLIPNLVYLGCORVLANKEIM 180
DB 121 LAGGFAPFRCFPGLECKGSTLVPTCISOPCLPVANIGPRLIPNLVYLGCORVLANKEIM 180
QY 181 QONGIGVLANASYTCPPKDFIPESHFLRVVNDSPCEKILPWLKSVDFTEKAKASNGCV 240
DB 181 QONGIGVLANASYTCPPKDFIPESHFLRVVNDSPCEKILPWLKSVDFTEKAKASNGCV 240
QY 241 LVHCLAGISRATIAIAYIKRMDSLDEAYRFEKERRPTISPNFNLGQLDYEKKIKX 300
DB 241 LVHCLAGISRATIAIAYIKRMDSLDEAYRFEKERRPTISPNFNLGQLDYEKKIKX 300
QY 301 QTGASGPKSLKLIHLEKNEPVPAVSEGGKSETPLSPCADSATSEAGORPVHPASV 360
DB 301 QTGASGPKSLKLIHLEKNEPVPAVSEGGKSETPLSPCADSATSEAGORPVHPASV 360
QY 361 PSVPVOPSLLEDSPVQALSGHLSDRLSDNKLKRSPLDIKSVSASMAASLHGF 420
DB 361 PSVPVOPSLLEDSPVQALSGHLSDRLSDNKLKRSPLDIKSVSASMAASLHGF 420
QY 421 SSEDLEVEYKPSSTLTDGNTKLCQFSPVCELSQETPETS-----PDKEBASIPKLIQTA 474
DB 421 SSEDLEVEYKPSSTLTDGNTKLCQFSPVCELSQETPETS-----PDKEBASIPKLIQTA 474
QY 475 RPEDSOSKR-----LHSVTSSTSGTQORSLSLPHRSGSVEDNHTYTFLEGLS 522
DB 475 RPEDSOSKR-----LHSVTSSTSGTQORSLSLPHRSGSVEDNHTYTFLEGLS 522
QY 523 TSQOHLTKSAGLGKMGHSDIILAPOTSPLTSSWYATSSHPYASATYGGASASVAY 582
DB 523 TSQOHLTKSAGLGKMGHSDIILAPOTSPLTSSWYATSSHPYASATYGGASASVAY 582
QY 583 SCQOLPTCGDQVYSVRRQRKPSDRADSRSMHESPEPEKQFKRRSCQMEFGESIMSENRS 642
DB 583 SCQOLPTCGDQVYSVRRQRKPSDRADSRSMHESPEPEKQFKRRSCQMEFGESIMSENRS 642
QY 643 REELGKVGSGSFPSSGMEIIEVS 665
DB 643 REELGKVGSGSFPSSGMEIIEVS 665
QY 644 REELGKVGSGSFPSSGMEIIEVS 666
DB 644 REELGKVGSGSFPSSGMEIIEVS 666

RESULT 3
AAM25744
ID AAM25744 standard; protein; 672 AA.

AC AAM25744;
XX 16-OCT-2001 (first entry)
DE Human protein sequence SEQ ID NO:1259.

XX Human; cancer; HIV infection; human immunodeficiency virus;
XX anti-inflammatory; antirheumatic; antiarthritic; immunosuppressive;
XX antibacterial; endocrine; cardiac; central nervous system; vitinide;
XX anti-HIV; fungicide; antitumor; cardiovascular; antiaesthetic; anemia;
XX antiagregant; haemostatic; vulnary; antileuc; osteopathic; eczema;
XX dermatological; antiallergic; antiaesthetic; antidiabetic; cyostatic;

KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunosuppressant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathic disorder; cardiac anaphylaxis; autoimmunity;
KW genetic diseases; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.

OS Homo sapiens.

PN WO200153455-A2.

PD 26-JUL-2001.

PF 22-DEC-2000; 2000MO-US035017.

PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

PI WPI: 2001-457603/49.

DR N-PSDB; AAM99685.

PT Isolated human polynucleotides encoding polypeptides, useful for the

PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.

PS Claim 20; Page 260; 1217pp; English.

XX AAM99166 to AAM99904 encode the human proteins given in AAM25225 to

XX AAM25963. The proteins can have activities based on the tissues and cells

XX they are expressed in, such as: anti-inflammatory; antirheumatic;

XX antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;

XX central nervous system; vitinide; anti-HIV; fungicide; antitumor;

XX cardiovascular; antiaesthetic; antileuc; haemostatic; vulnary;

XX antidiabetic; osteopathic; dermatological; antiallergic; antiaesthetic;

XX antiparkinsonian; and immunosuppressant. The proteins and polynucleotides

XX encoding them can be used in gene therapy, antisense therapy and vaccine

XX production. The proteins and polynucleotides are useful for screening for

XX agonists or antagonists of a protein and for the treatment and diagnosis of

XX disorders associated with the activity of a protein e.g. inflammation,

XX rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,

XX neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal

XX infections, autoimmunity, genetic diseases, haematopoietic disorders,

XX anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,

XX osteoporosis, severe combined immunodeficiency, eczema, allergic

XX rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,

XX Alzheimer's disease, Parkinson's disease, neurodegenerative and

SO Sequence 672 AA;

Query Match 85.7%; Score 2930; DB 1; Length 672;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 591; Conservative 10; Mismatches 46; Indels 36; Gaps 5;

QY 1 MAHEMIGTOIVTERLVALESGETEKVLLIDSRPFVEVNTSHLEAININCKMKRRLQ 60
DB 8 MAHEMIGTOIVTERLVALESGETEKVLLIDSRPFVEVNTSHLEAININCKMKRRLQ 67
QY 61 DKVLITELIQSHAKHVIDIDCQKVVVYDQSSQDVASLSDDCEFLVYLGLKLEKSPNSVHL 120
DB 61 DKVLITELIQSHAKHVIDIDCQKVVVYDQSSQDVASLSDDCEFLVYLGLKLEKSPNSVHL 120
QY 121 LAGGFAPFRCFPGLECKGSTLVPTCISOPCLPVANIGPRLIPNLVYLGCORVLANKEIM 180
DB 121 LAGGFAPFRCFPGLECKGSTLVPTCISOPCLPVANIGPRLIPNLVYLGCORVLANKEIM 187

Qy	18	QONGIGVLAASATCPKPDPIPESHILARVPANDSCCEKILPMLDKSVDFEIKAAASNCV	24.0
Db	188	QONGIGVLAASATCKEPPDPIPESHILARVPANDSCCEKILPMLDKSVDFEIKAAASNCV	24.7
Qy	241	LVMHLAGISHSATTAIAYIMKRMWMSLDEAYRFPYKGPPTISPNMFLGOLLDEYKIKN	30.0
Db	248	LVMHLAGISHSATTAIAYIMKRMWMSLDEAYRFPYKGPPTISPNMFLGOLLDEYKIKN	30.7
Qy	301	QTGASGKSKLTKLHLHEKNEPVPANVSEGGOKSETPLSPPCADSATSBAQRVHPASV	36.0
Db	308	QTGASGKSKLTKLHLHEKNEPVPANVSEGGOKSETPLSPPCADSATSBAQRVHPASV	36.7
Qy	361	PSVPSVQPSLLEDSPLVOALSGHLHSAEDLEDSNKLKRSFSLDIKSVSYSAASMAASHGF	42.0
Db	368	PSVPSVQPSLLEDSPLVOALSGHLHSAEDLEDSNKLKRSFSLDIKSVSYSAASMAASHGF	42.7
Qy	421	SSSEADLEYYKPESTTLDGNTKLCOQSPVOEL-----SEQTPENSPDKKSAISIKYLOTA	47.4
Db	428	SSSEADLEYYKPESTTLDGNTKLCOQSPVOEL-----SEQTPENSPDKKSAISIKYLOTA	48.7
Qy	475	RPSDGSOKR-----LHSVATSSSGTAQKSLPLHRSQSVEDNYHTSFLTGLS	52.2
Db	488	RASDCLRSEPAVAAPRGPFFYLHCIEVG---AWRTITP-----ASFSAFP	53.1
Qy	523	TSOOHLTKASGIGLKMHSDDILAPOTSFSLTSSMYKATETSSHYTSASAIYGSASISAY	56.2
Db	532	PAAPH--EYCMPEPKGLADPIIAPOTSTPSTLTSWYFATETSSHYTSASAIYGSASISAY	56.9
Qy	583	SCSQLPFCGDQVYSVARROKPSDRAISRSMHEESPFEKQPKRRSCQMEFEGESIMSHNR	64.2
Db	590	SCSQLPFCGDQVYSVARROKPSDRAISRSMHEESPFEKQPKRRSCQMEFEGESIMSHNR	64.9
Qy	643	REELGKVGSSSFSGSMELIIVS	66.5
Db	650	REELGKVGSSSFSGSMELIIVS	67.2

Search completed: July 1, 2004, 14:15:18
Job time : 1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:15:52 ; Search time 22.6939 Seconds

(without alignments)
2818.703 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 3418

Sequence: 1 MAHEMGTGIVTERLVALLD.....LGNVSGSSFSGSMETIEVS 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	497.5	14.6	619	2	T15969
2	443	13.0	394	2	A56115
3	433	12.7	367	1	S24411
4	423	12.4	367	2	S52265
5	420	12.3	367	1	S29090
6	418	12.2	314	1	A57126
7	409	12.0	314	2	B57126
8	393.5	11.5	393	2	A56947
9	360.5	10.5	384	1	I38890
10	339	9.9	384	1	T46405
11	294.5	8.6	365	2	T32494
12	239.5	7.0	272	2	T18915
13	233.5	6.8	226	2	T21810
14	225.5	6.6	186	2	T16056
15	223.5	6.5	330	2	T39698
16	208.5	6.1	220	2	JC7885
17	205.5	6.0	866	2	F88481
18	205	6.0	223	2	I49365
19	204	6.0	278	2	T39517
20	203.5	6.0	771	2	T47666
21	202	5.9	142	2	T03074
22	202	5.9	364	1	S31304
23	195	5.7	185	1	A47196
24	195	5.7	283	2	G84458
25	193	5.6	205	2	I49364
26	189	5.5	807	1	S58755
27	185	5.4	807	1	S44538
28	183.5	5.4	276	2	T48906
29	178	5.2	204	2	T17802

30	177	5.2	580	2	T18439	hypothetical prote
31	172	5.0	600	2	T18446	hypothetical prote
32	170.5	5.0	292	2	S41012	hypothetical prote
33	157	4.6	209	1	S48459	probable dual spec
34	150.5	4.4	169	2	T30684	probable dual spec
35	147.5	4.3	597	1	S43743	probable dual spec
36	146.5	4.3	272	2	T19418	hypothetical prote
37	145.5	4.3	171	1	I36845	dual specificity p
38	145.5	4.3	171	2	T28522	probable dual spec
39	145.5	4.3	171	2	B72161	UL protein - vari
40	144.5	4.2	171	1	Q0VZH1	dual specificity p
41	142.5	4.2	928	2	S50578	hypothetical prote
42	140	4.1	2271	2	F90073	hypothetical prote
43	139.5	4.1	534	2	T39903	serine-rich protei
44	138.5	4.1	171	1	A42514	dual specificity p
45	137.5	4.0	1906	1	S68235	myosin-light-chain

ALIGNMENTS

RESULT 1

T15969
hypothetical protein F08B1.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

R/Chromosome: S.

A/Description: The sequence of C. elegans cosmid F08B1.

A/Reference number: Z18439

A/Accession: T15969

A/Status: preliminary; translated from GB/EMBL/DDBT

A/Molecule type: DNA

A/Residues: 1-619 <CHI>

A/Cross-references: EMBL:U23178; NID:9726421; PID:9726422; PIDN:AAC6719.1; CESP:F08B1.1

A/Experimental source: strain Bristol N2

C/Genetics:

A/Gene: CESP:F08B1.1

A/Introns: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3

Query Match 14.6%; Score 497.5; DB 2; Length 619;
Best Local Similarity 28.9%; Pred. No. 4.4e-24;
Matches 160; Conservative 77; Mismatches 159; Indels 157; Gaps 19;

QY	123	GGFAERSRCFPGCEGKS--TLVPTCIQPC--PVANIGPTNILENLYGCGQDVLNKE	178
DB	99	GGFKQPAQGYPCQCSSEGWTRLPQSLQPCLSQPTGD-GITLITPVIYGSQIDSIDET	157
QY	179	LIQNGIGYVLANSYTCPRDPFPE-SHPLRVVNVNSFCCKILPWLKDSVDFTEKASN	237
DB	158	MLDALDISVYINISMTCPSKVCIKEDKFNIRIVNDSYQKSLSPYFMAVEFEKCRAG	217
QY	238	GCVLVHCLAGISRSATIAIYIMKRDMSLDEAYRFVKEKRPITSPFNFLGQLDYEK-	296
DB	218	KKCLHICLAGISRSPTLAIYIMKMGSDDAIRYKKEKRPISPFNFMGQLDYENV	277
QY	297	---KIKNQTAGSGPKSKLKL--HLEKNEP-----VPAVSEG	329
DB	278	LKIDVLDVYQAARPHRMDYVGPDLCPKVPKSSASNCVFGSTHDESSPSPVSEG	337
QY	330	GQKSETPPLSPCADSATSEAGRPVHPASVSPVSPQPSLDBSLVQALSGHLISADR	389
DB	338	SAASEPTSSAASSSTASNA-----PPMPSTSE-----QCTSSGTVAVNG	379
QY	390	LEDSNKLKSPFSLDI-----KSVYSASMAALHGFSSBDALEYKPTSLDGTNKL	443
DB	380	-----KKNMTMDLGLPHRPKALGLPRLGTSV-----ALPSPSTELSRLS---	420
QY	444	QSPVQVLESEQTP--ETSPDKKEASIPKQLQTPARPSDOSKALHVSRTSSGTAQSL	500
DB	421	-FNGPEALVAPSTILNFTNCPGNSPIIV-----ASSSRBVILTLTP	460

QY 501 SPLHRSQSVEDNHTSFLRGLSTSOOHLTKSAGLGLKGMHSDILAPQTSPELSTSSWYFA 560
 Db 461 TRAASSSS-----STS-----SEPFDDSSRES 483
 QY 561 TESSH-----FYASATYGGSAASYAVSCQLPTCGQGVSVRRQRKPSRADRRSMW 614
 Db 484 SSSSSIVENPFPASTEVPAGSSSISTPSGSG-----STPASASSAARSC 529
 QY 615 EESPEKOPKRRS 627
 Db 530 RMKGFFKVPFSKKA 542

RESULT 2

A56115

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 4 - human

N/Alternate names: dual specificity phosphatase HVH2

C/Species: Homo sapiens (man)

C/Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 24-Apr-1998

C/Accession: A56115

R/Guan, K.L.; Butch, E.

J. Biol. Chem. 270, 7197-7203, 1995

A/Title: Isolation and characterization of a novel dual specific phosphatase, HVH2, which

A/Reference number: A56115; MUID:95221370; PMID:7535768

A/Accession: A56115

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-394 <GUA>

A/Cross-references: GB:U21108

C/Genetics:

A/Gene: GDB:DUSP4; HVH2; MKP-2

A/Cross-references: GDB:433893

A/Map position: 8p21-8p11.2

C/Suprafamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity

C/Keywords: phosphoprotein, phosphoric monoester hydrolase

F/203-334/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>

F/280/Active site: Cys (phosphocysteine intermediate) #status predicted

F/286/Binding site: substrate phosphate (Arg) #status predicted

Query Match 13.0%; Score 443; DB 2; Length 394;
 Best Local Similarity 31.4%; Pred. No. 6.9e-21;
 Matches 120; Conservative 74; Mismatches 142; Indels 46; Gaps 14;

QY 19 LESGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQODKYLITELI--CHSAKHK 76
 Db 39 LPSSG-GKCLLDRCPLAHAGYILGSVNVCNTIVRR-AKGSVSLQGLIPAEVRRAR 96
 QY 77 VDIDCSQKVYVYDSSQDVASLSDCFITVLIGLTKSF--NSVHLLAGGFAEFSRCPPG 134
 Db 97 LRSGLYSAVIVYDERSPPAESLRDSTVSLVVOALRRNAERTDCLCKGGERFSSEYPE 156
 QY 135 LCEGKSTL-----VPTCISOP-----C-LPVANI-GPTRILPNTLYLGGQRDVNLKEL 179
 Db 157 FCSYTKMLAALPPVPBSATPELDLGCSSCSTPLHDGGPVEILLFYLIGSAHYAARDM 216
 QY 180 IQONGIGVYVNASYTCPKPDPFIPESHFLRPVNDSECEKILPMLDKSVDFIEKAKANGC 239
 Db 217 LDALGIVALLINVSDDCPN-HFEGHYQKCIPIVEDNHKADISSWMEALIEYIDAVKDCGR 275
 QY 240 VLVHCLAGISRSATITAYIMKRMDSIDEAYRFRKRPRTISNENFLGQLDYKKIK 299
 Db 276 VLVHCGAGISRSATITAYIMKRMKRVLEAEFVKORRSITISNFSFMQQLQFESQVL 335
 QY 300 NOTGASGPKSKLKLHLEKNEPVPVASEGQKSETPSPCADSATSEAGORVHPAS 359
 Db 336 ATSCAANAAS-----PSGPL-----REKKTATP-----TSQVFSPFPV-SVG 373
 QY 360 VPSVPSVQPSLLEDSPLVQALS 381
 Db 374 VHSAPSSLPYL--HSPITTPS 393

RESULT 3

S24411

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - mouse

N/Alternate names: 3CH134 protein; protein-tyrosine-phosphatase erp, nonreceptor type 10

C/Species: Mus musculus (house mouse)

C/Date: 19-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999

C/Accession: A54681; S24411

R/Nozuchi, T.; Metz, R.; Chen, L.; Mattei, M.G.; Carrasco, D.; Bravo, R.

Mol. Cell. Biol. 13, 5195-5205, 1993

A/Title: Structure, mapping, and expression of erp, a growth factor-inducible gene encodi

A/Reference number: A54681; MUID:93360956; PMID:8355678

A/Accession: A54681

A/Molecule type: DNA

A/Residues: 1-367 <NOG>

A/Cross-references: GB:S64851; NID:g409976; PIDN:AMB27882.1; PID:g409977

R/Charles, C.H.; Adler, A.S.; Lau, L.F.

Oncogene 7, 187-190, 1992

A/Title: cDNA sequence of a growth factor-inducible immediate early gene and characteriz

A/Reference number: S24411; MUID:92158357; PMID:1741163

A/Accession: S24411

A/Molecule type: mRNA

A/Residues: 1-367 <CHA>

A/Cross-references: EMBL:X61940; NID:g49735; PIDN:CAA43944.1; PID:g49736

C/Genetics:

A/Gene: erp

A/Introns: 123/1; 172/1; 245/1

C/Suprafamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity

C/Keywords: immediate-early protein; phosphoprotein; phosphoric monoester hydrolase

F/181-312/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>

F/258/Active site: Cys (phosphocysteine intermediate) #status predicted

F/264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.7%; Score 433; DB 1; Length 367;
 Best Local Similarity 31.6%; Pred. No. 2.7e-20;
 Matches 115; Conservative 54; Mismatches 151; Indels 44; Gaps 8;

QY 15 LVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQODKYLITELIOLHSAR 74
 Db 13 LRALLREGAQCLLDRCSPFAFNAGHILGSVNVRSTIVRRAKGAMGLEHIVPAEIR 72
 QY 75 HKVDIDCSQKVYVYDSSQDVASLSDCFITVLIGL--EKSNSVHLLAGGFAEFSRCPPG 132
 Db 73 GRLLAGAYHVAIVLLDRSASLIDGAKRDPGLALAAGLCREARSTQVFFLGGGEAFASAC 132
 QY 133 PGLCEGKSTLVPTCISOP-----CLPVANI-----GPTRIIPNTLYLGGQRDVNL 175
 Db 133 PELCSQST--PGLSLPLSTVPSDASGSCSTPLPDGGPVEILLFYLIGSAHYAS 190
 QY 176 NKELIQONGIGVYVNASYTCPKPDPFIPESHFLRPVNDSECEKILPMLDKSVDFIEKAKA 235
 Db 191 RKQMLDALGITALLINVSANCPN-HFEGHYQKSIPIVEDNHKADISSWMEALIDFIDSID 249
 QY 236 SNGCVVHCLAGISRSATITAYIMKRMDSIDEAYRFRKRPRTISNENFLGQLDYE 295
 Db 250 AGGRVVFHCGAGISRSATITAYIMKRMKRVLEAEFVKORRSITISNFSFMQQLQFESQVL 309
 QY 296 KRIKQGTGASGPKSKLKLHLEKNEPVPVASEGQKSETPSPCADSATSEAGORV 355
 Db 310 SQV-----LAHPCS-----AEAGSPMAVLDRGISTTVENFPVS-----IFV 347
 QY 356 HPAS 359
 Db 348 HPTN 351

RESULT 4

S52265 dual specificity phosphatase (EC 3.1.3.-) 1 - rat

N/Alternate names: protein-tyrosine-phosphatase C100; protein-tyrosine-phosphatase, non

C/Species: Rattus norvegicus (Norway rat)

C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 22-Jun-1999

C/Accession: S52265

R/Muda, M.; Schlegel, W.; Arkinsteall, S.

submitted to the EMBL Data Library, January 1995

A:Description: Pathways regulating CL100 gene expression in pituitary cells.
 A:Reference number: S52265
 A:Accession: S52265
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-367 <MUD>
 A:Cross-references: EMBL:X64004; NID:G642264; PIDN:CAA58828.1; PID:G642265
 C:Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
 C:Keywords: phosphoprotein; phosphoric monoester hydrolase
 F:181-312/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
 F:258/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.4%; Score 423; DB 2; Length 367;
 Best Local Similarity 31.3%; Pred. No. 1.2e-19;
 Matches 115; Conservative 54; Mismatches 148; Indels 50; Gaps 9;

15 LVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQ-----QDKVLTLELI 69
 13 LRLALGERAQCILLDCRSFFAFNAGHTGVSVVNRSTIYRRKAKGMGLHIVPAELR 71
 70 QHSKAKRVDIDCSQKVVVYDQSSQDVASLSDDCFVLTKL--EKSFNSVHLAAGFAE 127
 72 -----RGLLAGAVHAVVLLDERSAALDGAKRDDGLAAGALCREARSTQVFLGGYE 127
 128 FSRCPFGLECEK-----STLVPTCISQPCLPVANI-----GPTRIPLNLYGCGR 172
 128 FSAACGELCSKQSTPMGLSLPSTSVPSAEGSCSGSTPLVDGCGVEILSLYLSAY 187
 173 DVNLKELIQNGIGVYLNASTYCPKPDFIPESHFLRVVNDSPCEKILPWLKSVDFIEK 232
 188 HARKOMLADLGTITALINVSANCPN-HFEGHYQKSLIPEVDNKHADISSPFNEAIDPDS 246
 233 AKASNGCVLVHCLAGISRSATITAIYIMKMDSLDEAYRVEKRPITSPNFNLGOLL 292
 247 IKDAGGVFVHCQAGISRSATITAIYIMKMDSLDEAYRVEKRPITSPNFNLGOLL 306
 293 DYKAKTKNQTGASGPKSKLKLHLEKNEPVAVSEGGQKSET-----PLSPCADSATSEAG 352
 307 QFESQV-----LAFHCS-----AEAGSPMAVLDKRGSTITTVFNPVPS----- 344

QY 353 RPVHPAS 359
 DB 345 IYVHPIN 351

RESULT 5
 dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - human
 N:Alternate names: protein-tyrosine-phosphatase CL100; protein-tyrosine-phosphatase, not
 C:Species: Homo sapiens (man)
 C>Date: 25-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999
 C:Accession: S29090; A53052
 R:Keyes, S.M.; Emalle, E.A.
 Nature 359, 644-647, 1992
 A:Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine
 A:Reference number: S29090; MUID:93024952; PMID:1406996
 A:Accession: S29090
 A:Molecule type: mRNA
 A:Residues: 1-367 <KEY>
 A:Cross-references: EMBL:X68277; NID:G29980; PIDN:CAA6338.1; PID:G29981
 R:Wak, S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.
 J. Biol. Chem. 269, 3596-3604, 1994
 A:Title: Isolation and characterization of a human dual specificity protein-tyrosine pho
 A:Reference number: A53052; MUID:94148864; PMID:8106404
 A:Accession: A53052
 A:Molecule type: DNA
 A:Residues: 1-367 <KMA>
 A:Experimental source: leukocyte
 A:Note: Sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804,
 C:Genetics:
 A:Gene: GDB:DUSP1; PTPN10
 A:Cross-references: GDB:136197; OMIM:600714

A:Map position: 5q34-5q34
 C:Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
 C:Keywords: heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induced
 F:181-312/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
 F:258/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.3%; Score 420; DB 1; Length 367;
 Best Local Similarity 31.2%; Pred. No. 1.8e-19;
 Matches 113; Conservative 56; Mismatches 159; Indels 34; Gaps 7;

15 LVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQDDKVLTELIOHSK 74
 13 LRLALGERAQCILLDCRSFFAFNAGHTGVSVVNRSTIYRRKAKGMGLHIVPAELR 72
 75 HKYDIDCSQKVVVYDQSSQDVASLSDDCFVLTKL--EKSFNSVHLAAGFAEPRCF 132
 73 GRLLAGAVHAVVLLDERSAALDGAKRDDGLAAGALCREARAQVFLKGYEAFSASC 132
 133 PGLECEK-----STLVPTCISQPCLPVANI-----GPTRIPLNLYGCGRVLNK 177
 133 FELCSKQSTPMGLSLPSTSVPSAEGSCSGSTPLVDGCGVEILSLYLSAYASRK 192
 178 ELIQNGIGVYLNASTYCPKPDFIPESHFLRVVNDSPCEKILPWLKSVDFIEKASN 237
 193 DMIDLGITITALINVSANCPN-HFEGHYQKSLIPEVDNKHADISSPFNEAIDPDSIKNAG 251
 238 GCVLVHCLAGISRSATITAIYIMKMDSLDEAYRVEKRPITSPNFNLGOLLDEK 297
 252 GRVFNVCQAGISRSATITAIYIMKMDSLDEAYRVEKRPITSPNFNLGOLLDEK 311
 298 IKNTGASGPKSKLKLHLEKNEPVAVSEGGQKSET-----PLSPCADSATSEAG 353
 312 V-----LAFHCS-----AEAGSPMAVLDKRGSTITTVFNPVPSIPVHSTNSALSTLOS 359

QY 354 PV 355
 DB 360 PI 361

RESULT 6
 dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 2 - human
 N:Alternate names: mitogen-induced nuclear tyrosine phosphatase; phosphatase of activated
 C:Species: Homo sapiens (man)
 C>Date: 03-Nov-1995 #sequence_revision 02-Jul-1996 #text_change 11-Jun-1999
 C:Accession: A57126
 R:Rohan, P.J.; Davis, P.; Moskaluk, C.A.; Kearns, M.; Kruttsch, H.; Stiebenliet, U.; Kelly
 Science 259, 1763-1766, 1993
 A:Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
 A:Reference number: A57126; MUID:93206122; PMID:7681221
 A:Accession: A57126
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-314 <ROH>
 A:Cross-references: GB:111329; NID:G559539; PIDN:AAA50779.1; PID:G292376
 C:Genetics:
 A:Gene: GDB:DUSP2
 A:Cross-references: GDB:139200
 A:Map position: 2q11-2q11
 C:Function:
 A:Description: catalyzes the hydrolysis of peptidyl-phosphoserine, -phosphothreonine, and
 C:Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
 C:Keywords: nucleus; phosphoprotein; phosphoric monoester hydrolase
 F:140-160/Region: nuclear location signal
 F:180-311/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
 F:257/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:263/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.2%; Score 418; DB 1; Length 314;
 Best Local Similarity 36.0%; Pred. No. 2e-19;
 Matches 109; Conservative 45; Mismatches 105; Indels 44; Gaps 8;

C/Accession: T18915
 R/L10Yd, C.
 submitted to the EMBL Data Library, November 1996
 A/Reference number: Z19044
 A/Accession: T18915
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-272 <MTL>
 A/Cross-references: EMBL:Z81461; PIDN: CAB03837.1; GSPDB: GN00019; CESP: C04F12.8
 A/Experimental source: clone C04F12
 C/Genetics:
 A/Map position: 1
 A/Introns: 53/1; 84/3; 204/2

Query Match 7.0%; Score 239.5; DB 2; Length 272;
 Best Local Similarity 30.2%; Pred. No. 4, 2e-08;
 Matches 77; Conservative 45; Mismatches 110; Indels 23; Gaps 8;

OY 156 NIG----PTRLPLNYLGGCCRDVANKELIQONGIGYVLAASYCPKPDFIPESHFLRPV 211
 DB 4 NVGVLCQWSEINDHLYLS-GAGVLKPKDKIKQKIMIVNATTEPS-TYVQGVDTWKIRI 61
 OY 212 NDSECEKILPWLDSKVDPIEKAKASNGCVLVHCLAGISRSATIAIYIMKMDMSIDEAY 271
 DB 62 EDHFYARLNEHFDVADKINNVKRGKTLVHCHAGVSASLVMYLVHHEMTLRQAY 121
 OY 272 RFYKERPTISPNENFLGQLDYEEKIKNOTGASGPKSKLKLHLKEPNEPVASEG- 330
 DB 122 HYVAAPFTRPVNGVFWKQWVDEKRLRG-----TASVKWQVPECMPIFDVYADDI 174
 OY 331 --OKSEPTLSPCADSATSFAQGRVHPASV--PSVPSVQPSLEBSPLVQALSGHLISA 387
 DB 175 RMQINREISRHSLPFAQSSASQKRAFSASTRPSVSA--SSLATSSLRAYSPSSLP 232
 OY 388 DRLEDNKLRKSPSL 402
 DB 233 SSL----ALTTSYSL 243

RESULT 13

T21380
 hypothetical protein F26A3.4 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T21380
 R/McMurray, A.
 submitted to the EMBL Data Library, August 1996
 A/Reference number: Z19415
 A/Accession: T21380
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-226 <MTL>
 A/Cross-references: EMBL:Z78419; PIDN: CAB01700.1; GSPDB: GN00019; CESP: F26A3.4
 A/Experimental source: clone F26A3
 C/Genetics:
 A/Map position: 1
 A/Introns: 117/2; 150/3; 186/3

Query Match 6.8%; Score 233.5; DB 2; Length 226;
 Best Local Similarity 28.6%; Pred. No. 7, 7e-08;
 Matches 64; Conservative 47; Mismatches 80; Indels 33; Gaps 6;

OY 160 TRILPLNYLGGCCRDVANKELIQONGIGYVLAASYCPKPDFIPESHFLRPVNDSECEKI 219
 DB 15 SEIVPGFLI-CGVASLSEKDEKHKITHINATTEVPNLSLGDIGRTKLMELEDTPTQYI 73
 OY 220 LPMWDSKVDPIEKAKASNGCVLVHCLAGISRSATIAIYIMKMDMSIDEAYRFYKEKP 279
 DB 74 YPHLELQSDQIALIDGKVLVHCVAGVSRSSASISLAFLLKRCNLEAEVHAKSKS 133
 OY 280 TISNENFVLQQLDYEEKIKNOTGASGPKSKLKLHL-EPNEPVAVASEGQKSETPLS 338

DB 134 WVRPNLGFWRQLIAVEQVKNKAS-----VRLVREAGPEQLP----- 173
 OY 339 PPCADATSEAAQGRPVHPASVPSVQPSLEBSPLVQALSG 382
 DB 174 ----DVYINIAI---PARPAS---PRDDPMITPDEFRERRNSG 206

RESULT 14

T16056
 hypothetical protein F13D11.3 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001
 C/Accession: T16056
 R/Fulton, L.
 submitted to the EMBL Data Library, November 1995
 A/Description: The sequence of C. elegans cosmid F13D11.
 A/Reference number: S69020
 A/Accession: T16056
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-186 <FUL>
 A/Cross-references: EMBL:U40939; NID: g1072175; PID: g1072176; PIDN: AAA81700.1; CESP: F13D11
 C/Genetics:
 A/Map position: 30/3; 57/3; 85/3; 125/3; 172/3
 A/Introns: 30/3; 57/3; 85/3; 125/3; 172/3

Query Match 6.6%; Score 225.5; DB 2; Length 186;
 Best Local Similarity 32.9%; Pred. No. 1, 9e-07;
 Matches 57; Conservative 34; Mismatches 59; Indels 23; Gaps 5;

OY 160 TRILPLNYL---GQCRDVANKELIQONGIGYVLAASYCPKPDFIPESHFLRPVNDSEFC 216
 DB 12 TVRPHLFLAGYGC-----ITPSLKNQVITHGVDTNKTFR--IKGLDRLEVPVDVTL 65
 OY 217 EKLPMWDSKVDPIEKAKASNGCVLVHCLAGISRSATIAIYIMKMDMSIDEAYRFYKE 276
 DB 66 AKITQYFEPVIVYIEDAKQGHNTVYICAGVSSATILYLVWNTLSLEFAYLVQVQ 125
 OY 277 KRPTSPNENFLGQLDYEEKIKNOTGA---SGPKSKLKLHLKEPNEPVAV 326
 DB 126 VRFIISPNIGFWKQWVDEKRLRG-----PVPSV 167

RESULT 15

T39698
 protein tyrosine phosphatase - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C/Accession: T39698
 R/Hood, V.; Skelton, J.; Churche, C.M.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, July 1999
 A/Reference number: Z21870
 A/Accession: T39698
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-330 <MOO>
 A/Cross-references: EMBL:AL109652; PIDN: CAB51765.1; GSPDB: GN00067
 A/Experimental source: strain 972h-; cosmid c17A3
 C/Genetics:
 A/Map position: 2

Query Match 6.5%; Score 223.5; DB 2; Length 330;
 Best Local Similarity 39.3%; Pred. No. 5, 9e-07;
 Matches 53; Conservative 18; Mismatches 63; Indels 1; Gaps 1;

OY 160 TRILPLNYLGGCCRDVANKELIQONGIGYVLAASYCPKPDFIPESHFLRPVNDSECEKI 219
 DB 48 SEISKULYISSKWTABELVSTSKGIDYTLMSINPNLS-VPEQHLMLTIDESSQNTI 106
 OY 220 LPMWDSKVDPIEKAKASNGCVLVHCLAGISRSATIAIYIMKMDMSIDEAYRFYKEKP 279

Db	107	LOYEKSNNKFIAPALSKNAKVLVHCFAGISRSVTLVAAVLMKNNMTTEALSHINERRS	166
Qy	280	TISPENFPLGOLDY	294
Db	167	GISPNANPLRQLRKY	181

Search completed: June 21, 2004, 13:23:21
Job time : 23.6939 secs

This Page Blank (11/10/10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:15:22 ; Search time 52.2647 Seconds

(without alignments)
4014.554 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 3418
Sequence: 1 MAHEMIGCTIVTERLVALL.....LKVSGSQSPSGSMETIEVS 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3399	99.4	665	4	Q96N49	Q96N49 homo sapien
2	3379.5	98.9	662	4	Q81VT8	Q81VT8 homo sapien
3	3163	92.5	616	4	Q8N5T1	Q8N5T1 homo sapien
4	3079.5	90.1	660	11	Q920R2	Q920R2 mus musculus
5	2721.5	79.6	677	11	Q99MG6	Q99MG6 mus musculus
6	2380	69.6	622	11	Q99MG5	Q99MG5 mus musculus
7	1379.5	40.4	355	11	Q9AG16	Q9AG16 mus musculus
8	1327	38.8	625	4	Q86S88	Q86S88 homo sapien
9	1304	38.2	665	11	Q7TS29	Q7TS29 mus musculus
10	1038	30.4	300	11	Q9AG15	Q9AG15 mus musculus
11	891	26.1	206	11	Q8BZE4	Q8BZE4 mus musculus
12	611	17.9	657	5	Q8ST19	Q8ST19 caenorhabd
13	610	17.8	143	4	Q96GS2	Q96GS2 homo sapien
14	497.5	14.6	606	5	Q8ST18	Q8ST18 caenorhabd
15	487.5	14.3	367	13	Q7SZF3	Q7SZF3 brachydant
16	469	13.7	382	13	Q7J2L9	Q7J2L9 gallus gall

17	469	13.7	483	11	Q8R3L3	Q8R3L3 mus musculus
18	458.5	13.4	439	5	Q8IG35	Q8IG35 caenorhabd
19	449	13.1	383	13	Q8UW46	Q8UW46 fuju rubrip
20	445	13.0	368	4	Q8NFT0	Q8NFT0 homo sapien
21	445	13.0	411	4	Q13649	Q13649
22	439.5	12.9	398	11	Q8BFW3	Q8BFW3 mus musculus
23	437	12.8	382	13	Q7J2L8	Q7J2L8 brachydant
24	433.5	12.7	369	13	Q91790	Q91790 xenopus lae
25	431.5	12.6	369	13	Q90W58	Q90W58 xenopus lae
26	431	12.6	362	13	Q803B2	Q803B2 brachydant
27	427.5	12.5	378	13	Q916B3	Q916B3 xenopus lae
28	424	12.4	367	11	Q64193	Q64193 rattus norv
29	421	12.3	367	11	Q636B3	Q636B3 rattus norv
30	415.5	12.2	353	13	Q42253	Q42253 gallus gall
31	415	12.1	360	13	Q7ZVL8	Q7ZVL8 brachydant
32	413	12.1	318	11	Q80ZNL	Q80ZNL mus musculus
33	412	12.1	436	11	Q99KC2	Q99KC2 mus musculus
34	406	11.9	452	11	Q8K1S9	Q8K1S9 mus musculus
35	406	11.9	452	11	Q7TNL7	Q7TNL7 mus musculus
36	391.5	11.5	354	4	Q8N4A4	Q8N4A4 homo sapien
37	351.5	10.3	411	5	Q95SV1	Q95SV1 drosophila
38	351.5	10.3	411	5	Q9VW5	Q9VW5 drosophila
39	339	9.9	303	4	Q9NSW1	Q9NSW1 homo sapien
40	324.5	9.5	476	5	Q46122	Q46122 drosophila
41	319.5	9.3	476	5	Q9VW8	Q9VW8 drosophila
42	306	9.0	177	11	Q9CSL5	Q9CSL5 mus musculus
43	292.5	8.6	369	5	Q44128	Q44128 caenorhabd
44	287	8.4	1045	5	Q9NKT1	Q9NKT1 drosophila
45	287	8.4	1193	5	Q8IMU8	Q8IMU8 drosophila

ALIGNMENTS

RESULT 1
ID Q96N49 PRELIMINARY; PRT; 665 AA.
AC Q96N49;
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ1411.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuna M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Ohima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.,
RA "NEO human cDNA sequencing project."
RT Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
RL EMBL; AK055973; BAB71060.1; -.
DR HSSP; Q16828; IMR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:000470; F:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR00340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSGC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSGC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS0206; RHODANSE_3; 1.

DR PROSITE; PS00383; TYR PHOSPHATASE 1, 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein; Hydrolase.
SQ SEQUENCE 665 AA; 73058 MW; 1EAE0DF08460DF CRC64;

Query Match 99.4%; Score 3399; DB 4; Length 665;
Best Local Similarity 99.5%; Pred. No. 3.1e-252;
Matches 662; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMIGTOIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININGSKLMKRRLQ 60
DB 1 MAHEMIGTOIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININGSKLMKRRLQ 60
QY 61 DKVLITELIQAHSKHKVDIDCSQKVVYVDSSQVVASLSDCFITVLGLKLEKSFNSVHL 120
DB 61 DKVLITELIQAHSKHKVDIDCSQKVVYVDSSQVVASLSDCFITVLGLKLEKSFNSVHL 120
QY 121 LAGGFAPFRCPPGLCEGKSTLVPTCISQPCLPVANIPTILPNLYLGGORDVANKELI 180
DB 121 LAGGFAPFRCPPGLCEGKSTLVPTCISQPCLPVANIPTILPNLYLGGORDVANKELI 180
QY 181 QONGIGYVLNASYTTPKPDFIPESHFLRPVNDSECEKILPMLDKSVDFTEKAKANGCV 240
DB 181 QONGIGYVLNASYTTPKPDFIPESHFLRPVNDSECEKILPMLDKSVDFTEKAKANGCV 240
QY 241 LVHCLAGISRAATTAIATIMKMDMSLDEAYRFPYKERRPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRAATTAIATIMKMDMSLDEAYRFPYKERRPTISPNFNLGQLLDYEKKIKN 300
QY 301 QTASGPPSKKLHLHLEKNEPVAVSEGGQKSETPLSPCADSATSEAAQGPVHPASV 360
DB 301 QTASGPPSKKLHLHLEKNEPVAVSEGGQKSETPLSPCADSATSEAAQGPVHPASV 360
QY 361 PSVSVOPSLLEDSPLVQALSGHLSDRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
DB 361 PSVSVOPSLLEDSPLVQALSGHLSDRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
QY 421 SSSSDALEYKPPSTTLDGNTKLCQFSPVQELSEQTPETSPDKEBASIPKXQQTARPSDQ 480
DB 421 SSSSDALEYKPPSTTLDGNTKLCQFSPVQELSEQTPETSPDKEBASIPKXQQTARPSDQ 480
QY 481 SKRLHSVRTSSSGTAAORSLSLPLHRSQSVEDNYHTSFLFGISTSQOHLTKSAGLGLGWH 540
DB 481 SKRLHSVRTSSSGTAAORSLSLPLHRSQSVEDNYHTSFLFGISTSQOHLTKSAGLGLGWH 540
QY 541 SDILAQSTPSLTSSWYFATESSHFYASAIYGGASAYSVCQSLPTCGDQVYVRRR 600
DB 541 SDILAQSTPSLTSSWYFATESSHFYASAIYGGASAYSVCQSLPTCGDQVYVRRR 600
QY 601 OKPSRADRRRWHESSPEKQPKRRSCOMFEGSINSERNRRELKGVSGSSFSGSME 660
DB 601 OKPSRADRRRWHESSPEKQPKRRSCOMFEGSINSERNRRELKGVSGSSFSGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 2

QY 081VT8 PRELIMINARY; PRT; 662 AA.
DB 081VT8 PRELIMINARY; PRT; 662 AA.
AC 081VT8; PRT; 662 AA.
DT 01-MAR-2003 (Tremblrel_23, Created)
DT 01-MAR-2003 (Tremblrel_23, Last sequence update)
DE 01-OCT-2003 (Tremblrel_23, Last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eulheria; Primates; Catarrhini; Homidae; Homo.
OX NCB1_Taxid=9606;
RN
RP SEQUENCE FROM N.A.

RC TISSUE=Duodenum;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC042101; AAH42101.1; -.
DR GO; GO:0017017; P:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAP_kinase.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00782; DSCP; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHATASE.
DR SMART; SM00195; DSCP; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS0206; RHODANES 3; 1.
DR PROSITE; PS00383; TYR PHOSPHATASE 1; 1.
DR PROSITE; PS50056; TYR PHOSPHATASE 2; 1.
DR PROSITE; PS50054; TYR PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 662 AA; 72818 MW; 350534EF0652B98F CRC64;

Query Match 98.9%; Score 3379.5; DB 4; Length 662;
Best Local Similarity 99.2%; Pred. No. 9.6e-251;
Matches 660; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 MAHEMIGTOIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININGSKLMKRRLQ 60
DB 1 MAHEMIGTOIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININGSKLMKRRLQ 60
QY 61 DKVLITELIQAHSKHKVDIDCSQKVVYVDSSQVVASLSDCFITVLGLKLEKSFNSVHL 120
DB 61 DKVLITELIQAHSKHKVDIDCSQKVVYVDSSQVVASLSDCFITVLGLKLEKSFNSVHL 120
QY 121 LAGGFAPFRCPPGLCEGKSTLVPTCISQPCLPVANIPTILPNLYLGGORDVANKELI 180
DB 121 LAGGFAPFRCPPGLCEGKSTLVPTCISQPCLPVANIPTILPNLYLGGORDVANKELI 180
QY 181 QONGIGYVLNASYTTPKPDFIPESHFLRPVNDSECEKILPMLDKSVDFTEKAKANGCV 240
DB 181 QONGIGYVLNASYTTPKPDFIPESHFLRPVNDSECEKILPMLDKSVDFTEKAKANGCV 240
QY 241 LVHCLAGISRAATTAIATIMKMDMSLDEAYRFPYKERRPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRAATTAIATIMKMDMSLDEAYRFPYKERRPTISPNFNLGQLLDYEKKIKN 300
QY 301 QTASGPPSKKLHLHLEKNEPVAVSEGGQKSETPLSPCADSATSEAAQGPVHPASV 360
DB 301 QTASGPPSKKLHLHLEKNEPVAVSEGGQKSETPLSPCADSATSEAAQGPVHPASV 360
QY 361 PSVSVOPSLLEDSPLVQALSGHLSDRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
DB 361 PSVSVOPSLLEDSPLVQALSGHLSDRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
QY 421 SSSSDALEYKPPSTTLDGNTKLCQFSPVQELSEQTPETSPDKEBASIPKXQQTARPSDQ 480
DB 421 SSSSDALEYKPPSTTLDGNTKLCQFSPVQELSEQTPETSPDKEBASIPKXQQTARPSDQ 480
QY 481 SKRLHSVRTSSSGTAAORSLSLPLHRSQSVEDNYHTSFLFGISTSQOHLTKSAGLGLGWH 540
DB 481 SKRLHSVRTSSSGTAAORSLSLPLHRSQSVEDNYHTSFLFGISTSQOHLTKSAGLGLGWH 540
QY 541 SDILAQSTPSLTSSWYFATESSHFYASAIYGGASAYSVCQSLPTCGDQVYVRRR 600
DB 541 SDILAQSTPSLTSSWYFATESSHFYASAIYGGASAYSVCQSLPTCGDQVYVRRR 600
QY 601 OKPSRADRRRWHESSPEKQPKRRSCOMFEGSINSERNRRELKGVSGSSFSGSME 660
DB 601 OKPSRADRRRWHESSPEKQPKRRSCOMFEGSINSERNRRELKGVSGSSFSGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 3
ID Q8N5T1 PRELIMINARY; PRT; 616 AA.
AC Q8N5T1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Lung;
RA Straubberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC031643; AAI31643.1; -
DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro: IPR000343; MAPK_phosph.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSpC; 1.
DR PRINTS, PRO1764; MAPKPHPTASE.
DR SMART; SM00195; DSpC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 616 AA; 67636 MW; 2CBOB14482F2AD72 CRC64;
Query Match 92.5%; Score 3163; DB 4; Length 616;
Best Local Similarity 99.5%; Pred. No. 3.7e-234;
Matches 613; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 50 CSEKLMKRLQODKYLITELIQHSAKHKVDIDCSQKVVVYQSSQDVASLSDDCLTYLLG 109
DB 1 CSEKLMKRLQODKYLITELIQHSAKHKVDIDCSQKVVVYQSSQDVASLSDDCLTYLLG 60
QY 110 KLEKSFNSVHLLGGFAEFRCFPGLCCEKSTLVPTCISOPCLPVANIGFTRILPNLYLG 169
DB 61 KLEKSFNSVHLLGGFAEFRCFPGLCCEKSTLVPTCISOPCLPVANIGFTRILPNLYLG 120
QY 170 QORDVNLKELIQNGIGYVNAASYTCRPPDIPESHFLRPVNDSPCEKILPMLDKSVDF 229
DB 121 QORDVNLKELIQNGIGYVNAASYTCRPPDIPESHFLRPVNDSPCEKILPMLDKSVDF 180
QY 230 IERAKASNGCVLVHCLAGISRSATIAIYIMKMDMSLDEAYRFPVKEKRPITISPNENFLG 289
DB 181 IERAKASNGCVLVHCLAGISRSATIAIYIMKMDMSLDEAYRFPVKEKRPITISPNENFLG 240
QY 290 QILDYEKKINQNGASGPKSKLKLHLEKNEPVPANSEGGQSEFPLSPCCADSATSEA 349
DB 241 QILDYEKKINQNGASGPKSKLKLHLEKNEPVPANSEGGQSEFPLSPCCADSATSEA 300
QY 350 AGORPVHAPVSPVSPQSLLEDSPVQALSGHLHSADRLSDSNLKRFSFLDIKSVSY 409
DB 301 AGORPVHAPVSPVSPQSLLEDSPVQALSGHLHSADRLSDSNLKRFSFLDIKSVSY 360
QY 410 SASMAASLHGFSSEEDALEYKQSTTLDGTNKLQFSPVQELSBQIPETISDYEBASIPK 469
DB 361 SASMAASLHGFSSEEDALEYKQSTTLDGTNKLQFSPVQELSBQIPETISDYEBASIPK 420
QY 470 KLTOTARSDSOSKRLHSVRTSSSGTAORSLSPLRHSQVEDNPHYSFLGLSTSOQHLT 529
DB 421 KLTOTARSDSOSKRLHSVRTSSSGTAORSLSPLRHSQVEDNPHYSFLGLSTSOQHLT 480
QY 530 KSAAGLGLGKMHSDILAPOTSTPSLTSSGWYPATESSHFYASAIYGGASYSAYSCSOLPT 589
DB 481 KSAAGLGLGKMHSDILAPOTSTPSLTSSGWYPATESSHFYASAIYGGASYSAYSCSOLPT 540

QY 590 CGDQVYSVRRROKSPDRADSRBSWHESEPEKQFKRSCOMERGESIMSENRREELGKY 649
DB 541 CGDQVYSVRRROKSPDRADSRBSWHESEPEKQFKRSCOMERGESIMSENRREELGKY 600
QY 650 GSQSFSGSMELIEVS 665
DB 601 GSQSFSGSMELIEVS 616
RESULT 4
ID Q920R2 PRELIMINARY; PRT; 660 AA.
AC Q920R2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MAP kinase phosphatase-7.
GN DUSP16 OR 3830417M17RIK OR MKP-7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
RT "MKP-7, a Novel Mitogen-activated Protein Kinase Phosphatase,
RT Functions as a Shuttle Protein."
RL J. Biol. Chem. 276:39002-39011(2001).
DR EMBL: AB052157; BAB47240.1; -
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1917936; Dusp16.
DR GO: GO:0005737; C:cyclopasme; IDA.
DR GO: GO:0005634; C:nucleus; IDA.
DR GO: GO:000515; P:protein binding; IPI.
DR GO: GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSpC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS, PRO1764; MAPKPHPTASE.
DR SMART; SM00195; DSpC; 1.
DR PROSITE; PS50206; RHODANESE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Kinase.
SQ SEQUENCE 660 AA; 72695 MW; DB609FCDAD4AA309 CRC64;
Query Match 90.1%; Score 3079.5; DB 11; Length 660;
Best Local Similarity 90.1%; Pred. No. 1.1e-227;
Matches 599; Conservative 28; Mismatches 33; Indels 5; Gaps 3;
QY 1 MAHEMIGTOIVTERLVALLSGTEKYLIDSRPFVEYNTSHILEAININCKLMKRLQ 60
DB 1 MAHEMIGTOIVTERLVALLSGTEKYLIDSRPFVEYNTSHILEAININCKLMKRLQ 60
QY 61 DKYLITELIQHSAKHKVDIDCSQKVVVYQSSQDVASLS9DCFLTYLLGLKLEKSFNSVH 120
DB 61 DKYLITELIQHSAKHKVDIDCSQKVVVYQSSQDVASLS9DCFLTYLLGLKLEKSFNSVH 120
QY 121 LAGGFAEFRCFPGLCCEKSTLVPTCISOPCLPVANIGFTRILPNLYLGQORDVNLKELI 180
DB 121 LAGGFAEFRCFPGLCCEKSTLVPTCISOPCLPVANIGFTRILPNLYLGQORDVNLKELI 180
QY 181 QQNGIGYVNAASYTCRPPDIPESHFLRPVNDSPCEKILPMLDKSVDFIERAKASNGCV 240
DB 181 QQNGIGYVNAASYTCRPPDIPESHFLRPVNDSPCEKILPMLDKSVDFIERAKASNGCV 240

```

QY 241 LVHCLAGISRSATATAYIMKMDMSLDEAYRFVKEKPTTISPNFNFGQLLDYEKTIKN 300
Db 241 LVHCLAGISRSATATAYIMKMDMSLDEAYRFVKEKPTTISPNFNFGQLLDYEKTIKN 300
QY 301 QTASGPKSKLKLHLHLEKNEPVPAVSEGGOKSETPLESPCADSATSEAGORPVHPASV 360
Db 301 QTASGPKSKLKLHLHLEKNEPVPAVSEGGOKSETPLESPCADSATSEAGORPVHPASV 360
QY 361 PSVPSVQPSLLEDSPLVQALSGMLH.SADRLSDSNLKLKSPSLDIKVSYSASMAASLHGF 420
Db 361 PSVPSVQPSLLEDSPLVQALSGMLH.SADRLSDSNLKLKSPSLDIKVSYSASMAASLHGF 420
QY 421 SSEDALFYKPSSTLIDGTNKLQFSPVOELSEQTPETSPDKEASIPKQLQTPRPDSQ 480
Db 421 SSEDALFYKPSSTLIDGTNKLQFSPVOELSEQTPETSPDKEASIPKQLQTPRPDSQ 480
QY 481 SKRLHSVTRTSSGKTQORSLSPILHRSSEVENNYHTSLFGLSTSQOHLTKSAGLGLKGMH 540
Db 481 SKRLHSVTRTSSGKTQORSLSPILHRSSEVENNYHTSLFGLSTSQOHLTKSAGLGLKGMH 540
QY 541 SDIILAPQSTSTSLTSSWTFATESSHFYASAIYGSASYSAGSCQLPTCGDQVYSVRR 600
Db 541 SDIILAPQSTSTSLTSSWTFATESSHFYASAIYGSASYSAGSCQLPTCGDQVYSVRR 600
QY 601 QKPSDRADSRSRMSESPFEKQFRRCSCOMEFGESIMSENRSRRELGKVSQSSFGSGME 660
Db 601 QKPSDRADSRSRMSESPFEKQFRRCSCOMEFGESIMSENRSRRELGKVSQSSFGSGME 660
QY 661 IIEVS 665
Db 661 IIEVS 665

```

RESULT 5

```

QY 099MG6 PRELIMINARY; PRT; 677 AA.
Db 099MG6 PRELIMINARY; PRT; 677 AA.
QY 01-JUN-2001 (TREMBlrel. 17, Created)
Db 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
QY 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
Db 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
QY Map kinase phosphatase-M A1 isoform.
Db Map kinase phosphatase-M A1 isoform.
QY Mus musculus (Mouse).
Db Mus musculus (Mouse).
QY Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Db Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
QY Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Db Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
QY NCBI_TaxID=10090;
Db NCBI_TaxID=10090;
QY [1] SEQUENCE FROM N.A.
Db [1] SEQUENCE FROM N.A.
QY RC STRAIN=BALB/c;
Db RC STRAIN=BALB/c;
QY Matsumuchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
Db Matsumuchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
QY "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
Db "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
QY Activation in Macrophages."
Db Activation in Macrophages."
QY Mol. Cell. Biol. 20:6999-7009(2001).
Db Mol. Cell. Biol. 20:6999-7009(2001).
QY EMBL; AF345951; AAK35052.1; -.
Db EMBL; AF345951; AAK35052.1; -.
QY HSP; 016828; IMRP.
Db HSP; 016828; IMRP.
QY MGD; MGI:1917936; Dusp16.
Db MGD; MGI:1917936; Dusp16.
QY GO; GO:0005737; Cytoplasm; IDA.
Db GO; GO:0005737; Cytoplasm; IDA.
QY GO; GO:0005634; Cytoplasm; IDA.
Db GO; GO:0005634; Cytoplasm; IDA.
QY GO; GO:0005151; F:protein binding; IPI.
Db GO; GO:0005151; F:protein binding; IPI.
QY InterPro; IPR000340; DS phosphatase.
Db InterPro; IPR000340; DS phosphatase.
QY InterPro; IPR008343; MAPK_Chopsh.
Db InterPro; IPR008343; MAPK_Chopsh.
QY InterPro; IPR001763; Rhodanese-like.
Db InterPro; IPR001763; Rhodanese-like.
QY Pfam; PF00782; DSPC; 1.
Db Pfam; PF00782; DSPC; 1.
QY Pfam; PF00581; Rhodanese; 1.
Db Pfam; PF00581; Rhodanese; 1.
QY SMART; SM00195; DSPC; 1.
Db SMART; SM00195; DSPC; 1.
QY SMART; SM00450; RHOD; 1.
Db SMART; SM00450; RHOD; 1.
QY PROSITE; PS50206; RHODANASE_3; 1.
Db PROSITE; PS50206; RHODANASE_3; 1.
QY PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
Db PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.

```

```

DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
QY PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
Db PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
QY KW Hydrophobic; Kinase.
Db KW Hydrophobic; Kinase.
QY SEQUENCE 677 AA; 74550 MW; 886D5B7096C8C2FC CRC64;
Db SEQUENCE 677 AA; 74550 MW; 886D5B7096C8C2FC CRC64;

```

```

Query Match 79.6%; Score 2721.5; DB 11; Length 677;
Best Local Similarity 84.4%; Pred. No. 3.5e-200;
Matches 531; Conservative 36; Mismatches 57; Indels 5; Gaps 3;

```

```

QY 1 MAHEMIGQIVTERIVALLESSTKVLIDSRPVEVNTSHILEAININCSKLMKRLQ 60
Db 1 MAHEMIGQIVTERIVALLESSTKVLIDSRPVEVNTSHILEAININCSKLMKRLQ 60
QY 61 DKVITTELQHSAGHKVDICQSVVYDDSSQDVASLSDCHLYVLGLKESFNSVHL 120
Db 61 DKVITTELQHSAGHKVDICQSVVYDDSSQDVASLSDCHLYVLGLKESFNSVHL 120
QY 61 DVLITTELQHSAGHKVDICQSVVYDDSSQDVASLSDCHLYVLGLKESFNSVHL 120
Db 61 DVLITTELQHSAGHKVDICQSVVYDDSSQDVASLSDCHLYVLGLKESFNSVHL 120
QY 121 LAGGFAEFSRCFPCGCKSTLVPCTCISQPCLPVANIQTPLILNLYLGLCQDVLNKL 180
Db 121 LAGGFAEFSRCFPCGCKSTLVPCTCISQPCLPVANIQTPLILNLYLGLCQDVLNKL 180
QY 181 QONGIGVILNASTCPKDPFIPESHFLRVVNDSCFKILPWLKSVDFIEKAKASNGCV 240
Db 181 QONGIGVILNASTCPKDPFIPESHFLRVVNDSCFKILPWLKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATATAYIMKMDMSLDEAYRFVKEKPTTISPNFNFGQLLDYEKTIKN 300
Db 241 LVHCLAGISRSATATAYIMKMDMSLDEAYRFVKEKPTTISPNFNFGQLLDYEKTIKN 300
QY 301 QTASGPKSKLKLHLHLEKNEPVPAVSEGGOKSETPLESPCADSATSEAGORPVHPASV 360
Db 301 QTASGPKSKLKLHLHLEKNEPVPAVSEGGOKSETPLESPCADSATSEAGORPVHPASV 360
QY 361 PSVPSVQPSLLEDSPLVQALSGMLH.SADRLSDSNLKLKSPSLDIKVSYSASMAASLHGF 420
Db 361 PSVPSVQPSLLEDSPLVQALSGMLH.SADRLSDSNLKLKSPSLDIKVSYSASMAASLHGF 420
QY 421 SSEDALFYKPSSTLIDGTNKLQFSPVOELSEQTPETSPDKEASIPKQLQTPRPDSQ 480
Db 421 SSEDALFYKPSSTLIDGTNKLQFSPVOELSEQTPETSPDKEASIPKQLQTPRPDSQ 480
QY 481 SKRLHSVTRTSSGKTQORSLSPILHRSSEVENNYHTSLFGLSTSQOHLTKSAGLGLKGMH 540
Db 481 SKRLHSVTRTSSGKTQORSLSPILHRSSEVENNYHTSLFGLSTSQOHLTKSAGLGLKGMH 540
QY 541 SDIILAPQSTSTSLTSSWTFATESSHFYASAIYGSASYSAGSCQLPTCGDQVYSVRR 600
Db 541 SDIILAPQSTSTSLTSSWTFATESSHFYASAIYGSASYSAGSCQLPTCGDQVYSVRR 600
QY 601 QKPSDRADSRSRMSESPFEKQFRRCSCOMEFGESIMSENRSRRELGKVSQSSFGSGME 660
Db 601 QKPSDRADSRSRMSESPFEKQFRRCSCOMEFGESIMSENRSRRELGKVSQSSFGSGME 660
QY 661 IIEVS 665
Db 661 IIEVS 665

```

RESULT 6

```

QY 099MG5 PRELIMINARY; PRT; 622 AA.
Db 099MG5 PRELIMINARY; PRT; 622 AA.
QY 01-JUN-2001 (TREMBlrel. 17, Created)
Db 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
QY 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
Db 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
QY Map kinase phosphatase-M A2 isoform.
Db Map kinase phosphatase-M A2 isoform.
QY Mus musculus (Mouse).
Db Mus musculus (Mouse).
QY Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Db Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
QY Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Db Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
QY NCBI_TaxID=10090;
Db NCBI_TaxID=10090;
QY [1] SEQUENCE FROM N.A.
Db [1] SEQUENCE FROM N.A.
QY RC STRAIN=BALB/c;
Db RC STRAIN=BALB/c;
QY Matsumuchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
Db Matsumuchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
QY "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
Db "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
QY Activation in Macrophages."
Db Activation in Macrophages."
QY Mol. Cell. Biol. 20:6999-7009(2001).
Db Mol. Cell. Biol. 20:6999-7009(2001).
QY EMBL; AF345951; AAK35052.1; -.
Db EMBL; AF345951; AAK35052.1; -.
QY HSP; 016828; IMRP.
Db HSP; 016828; IMRP.
QY MGD; MGI:1917936; Dusp16.
Db MGD; MGI:1917936; Dusp16.
QY GO; GO:0005737; Cytoplasm; IDA.
Db GO; GO:0005737; Cytoplasm; IDA.
QY GO; GO:0005634; Cytoplasm; IDA.
Db GO; GO:0005634; Cytoplasm; IDA.
QY GO; GO:0005151; F:protein binding; IPI.
Db GO; GO:0005151; F:protein binding; IPI.
QY InterPro; IPR000340; DS phosphatase.
Db InterPro; IPR000340; DS phosphatase.
QY InterPro; IPR008343; MAPK_Chopsh.
Db InterPro; IPR008343; MAPK_Chopsh.
QY InterPro; IPR001763; Rhodanese-like.
Db InterPro; IPR001763; Rhodanese-like.
QY Pfam; PF00782; DSPC; 1.
Db Pfam; PF00782; DSPC; 1.
QY Pfam; PF00581; Rhodanese; 1.
Db Pfam; PF00581; Rhodanese; 1.
QY SMART; SM00195; DSPC; 1.
Db SMART; SM00195; DSPC; 1.
QY SMART; SM00450; RHOD; 1.
Db SMART; SM00450; RHOD; 1.
QY PROSITE; PS50206; RHODANASE_3; 1.
Db PROSITE; PS50206; RHODANASE_3; 1.
QY PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
Db PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.

```

RT Activation in Macrophages."
 RL Mol. Cell. Biol. 20:699-709(2001).
 DR EMBL; AF345952; AAK35053.1; -.
 DR HSSP; Q16828; IMKP.
 DR MGD; MGI:1917936; Duap16.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:000188; P:inactivation of MAPK; IDA.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PRO1764; MAPKPHPTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANSE_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR HydroLase; Kinase.
 SQ SEQUENCE 622 AA; 68672 MW; 6C0CA4E6E909B98 CRC64;

Query Match 69.6%; Score 2380; DB 11; Length 622;
 Best Local Similarity 75.7%; Pred. No. 5.2e-174;
 Matches 476; Conservative 36; Mismatches 57; Indels 60; Gaps 4;

QY 1 MAHEMIGTQVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 DB 1 MAHEMIGTQVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 QY 61 DKVLTIELIQAHAHAKVDIDCSQKVVYDSSQDVASLSDDCELYTLGLKESFNSVHL 120
 DB 61 DKVLTIELIQAHAHAKVDIDCSQKVVYDSSQDVASLSDDCELYTLGLKESFNSVHL 120
 QY 61 DKVLTIELIQAHAHAKVDIDCSQKVVYDSSQDVASLSDDCELYTLGLKESFNSVHL 120
 DB 61 DKVLTIELIQAHAHAKVDIDCSQKVVYDSSQDVASLSDDCELYTLGLKESFNSVHL 120
 QY 121 LAGFAEFRCFPGCEGKSTLVPTCISQPCLVANIGPRILPNLYLGCQDVLNKELI 180
 DB 121 LA-----DLM 125
 QY 181 QONGIGVYLAASYCPKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 DB 181 QONGIGVYLAASYCPKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 QY 126 QONGIGVYLAASYCPKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 185
 DB 126 QONGIGVYLAASYCPKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 185
 QY 241 LVHCLAGISRASATIAIAYIMKMDMSIDEAYRFVKEKRPITISPNFNLGQLDYEKKIKN 300
 DB 186 LVHCLAGISRASATIAIAYIMKMDMSIDEAYRFVKEKRPITISPNFNLGQLDYEKKIKN 245
 QY 301 QRTASGPKSLKLIHLEKNEPVPVASEGQKSETPSPCCADSATSEAGQRPVPVASY 360
 DB 246 QRTASGPKSLKLIHLEKNEPVPVASEGQKSETPSPCCADSATSEAGQRPVPVASY 302
 QY 361 PSVPSPQSLLEPDSPLYQALSGHLASADRLSDNKLKRSFLDIKSYASASMAASLHG 420
 DB 303 -SVPSLQPSLLEPDSPLYQALSGHLASADRLSDNKLKRSFLDIKSYASASMAASLHG 361
 QY 421 SSSSEDALEYYKPSPTLIDGTNKLQFSPVQVSESPETSPDKKEASIPKKLQTPRPSDQ 480
 DB 362 -SSSEDALEYYKPSPTLIDGTNKLQFSPVQVSESPETSPDKKEASIPKKLQTPRPSDQ 420
 QY 481 SKRLHVSRTSSSTTAQRSLSPHRSQSVNDNHTSFLRGLSTSQOHLTKSAGLGLKGM 540
 DB 421 VTLHLSVRTSSSTTAQRSLSPHRSQSVNDNHTSFLRGLSTSQOHLTKSAGLGLKGM 480
 QY 541 SDILAPQTSPLTSSWYFATSSSHFYASAGIYGSASAYASCQULPTCGDOVYSVRR 600
 DB 481 SDILAPQTSPLTSSWYFATSSSHFYASAGIYGSASAYASCQULPTCGDOVYSVRR 540
 QY 601 QKPSDRAHSRSHWESPFQKFRSSQ 629
 DB 541 QKPTDRADSRRTGMKRAPLKSSLNABAAK 569

RESULT 7

ID 09AG16 PRELIMINARY; PRT; 355 AA.
 AC 09AG16;
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Map Kinase phosphatase-M B1 isoform.
 GN DUSP16 OR 3830417M17RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
 RA Yoshikai Y.;
 RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
 RT Activation in Macrophages."
 RL Mol. Cell. Biol. 20:699-709(2001).
 DR EMBL; AF345953; AAK35054.1; -.
 DR HSSP; Q16828; IMKP.
 DR MGD; MGI:1917936; Duap16.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:000188; P:inactivation of MAPK; IDA.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PRO1764; MAPKPHPTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANSE_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR HydroLase; Kinase.
 SQ SEQUENCE 355 AA; 39502 MW; D52C29AE215CA285 CRC64;

Query Match 40.4%; Score 1379.5; DB 11; Length 355;
 Best Local Similarity 83.2%; Pred. No. 1.5e-97;
 Matches 273; Conservative 13; Mismatches 23; Indels 19; Gaps 2;

QY 1 MAHEMIGTQVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 DB 1 MAHEMIGTQVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
 QY 61 DKVLTIELIQAHAHAKVDIDCSQKVVYDSSQDVASLSDDCELYTLGLKESFNSVHL 120
 DB 61 DKVLTIELIQAHAHAKVDIDCSQKVVYDSSQDVASLSDDCELYTLGLKESFNSVHL 120
 QY 121 LAGFAEFRCFPGCEGKSTLVPTCISQPCLVANIGPRILPNLYLGCQDVLNKELI 180
 DB 121 LAGFAEFRCFPGCEGKSTLVPTCISQPCLVANIGPRILPNLYLGCQDVLNKELI 180
 QY 181 QONGIGVYLAASYCPKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 DB 181 QONGIGVYLAASYCPKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 QY 241 LVHCLAGISRASATIAIAYIMKMDMSIDEAYRFVKE-----KRPITSPNFNLG 289
 DB 241 LVHCLAGISRASATIAIAYIMKMDMSIDEAYRFVKE-----KRPITSPNFNLG 289
 QY 290 QLLDYEKKIKNGT-----ASGPKS 309
 DB 301 AKMNLERALCRRTGPRSWARWASBPAS 328

RESULT 8

Q86SS8

PRELIMINARY;

PRT; 625 AA.

AC Q86SS8
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-UN-2003 (Tremblrel. 25, Last sequence update)
 DT 01-0CT-2003 (Tremblrel. 25, Last annotation update)
 DE Similar to dual specificity phosphatase 8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, and Astrocytoma;
 RA Strauberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC045110; AAH45110.1; -
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR GO; GO:0006470; P:MAP kinase phosphatase activity; IEA.
 DR InterPro; IPR000340; DS phosphatase.
 DR InterPro; IPR008343; MAPK phosph.
 DR InterPro; IPR002965; P rich extensm.
 DR InterPro; IPR001763; Rhodanese-like.
 DR Pfam; PF00782; DSpC, 1.
 DR PRINTS; PRO1764; MAPKPHPTASE.
 DR PRINTS; PRO1217; PRICHEXTENS.
 DR SMART; SM00450; RHOD; 1.
 DR SMART; SM00195; DSpC; 1.
 DR PROSITE; PS50206; RHODANSE 3; 1.
 DR PROSITE; PS00383; TYR PHOSPHATASE 1; 1.
 DR PROSITE; PS50056; TYR PHOSPHATASE 2; 1.
 DR PROSITE; PS50054; TYR PHOSPHATASE DUAL; 1.
 DR SEQUENCE 625 AA; 65826 MW; C7C808407B724FFC CRC64;

Query Match

Best Local Similarity 38.8%; Score 1327; DB 4; Length 625;
 Matches 314; Conservative 94; Mismatches 185; Indels 104; Gaps 21;

1 MAHEMIGTOIV-TERLVALLESGETEKVLLIDSRPFVEYNTSHLEAININCSKIMKRLQ 59
 1 MAGDRLPKRYMDAKKLASLRGPGGPIVYDSRFSVEYNSMHWLVSVMCCSKLVKRLQ 60
 60 QDKVITLIELIQAAXKHVDIDCSQKVVYVYDSSQVAVLSDCFITVLGLKLEKSFNSVH 119
 61 QOKVITLIELIQAAXKHVDIDCSQKVVYVYDSSQVAVLSDCFITVLGLKLEKSFNSVH 119
 120 LIAGEPAEFSRCFELCEGK-STVPTCISQPCUVPANIGTRITLPMNYLGGCORVLTNKE 178
 121 ILTGFATFSSCFELCEGKPAALLPMGLSPCLPVPSPVGLTRILPHLYLGSQKDVANKD 180
 179 LIOGNGIGVYVNAATYCKPDPPEIPESHRLRPVNDSCFEKLLPWLKDSVDFTEKAKXNSG 238
 181 LMTQNGISYVNAASCPKDPFICBSRPMRPVINDNCEKLLPWLKDSIFIDKAKLSSC 240
 239 CVLVHCLAGISRSATITAIYIMKMDSIDEAFFVEKEKPTTSPNNFGLGOLLDYEKTI 298
 241 QVIVHCLAGISRSATITAIYIMKMDSIDEAFFVEKEKPTTSPNNFGLGOLLDYEKTI 300
 299 KNOGAGAGPKSKKLHLLEKPN---EPVPAVSEGGQKSETPPLSPCCADSNAT--SEAAGOR 353
 301 KLTLAALQDDPG-----TPSGTPPEPPSPPAAGAPLPRLP--PPTSESAATGNAALARG 350
 354 PVHPSVPSVPSVPSLLEDSPLVQALSGHLASDRLENSKLRKPSLIDIKSVISASAM 413
 351 GLASGGRPAAPPRPA---TSAQOGLRGHLHSSDRLODNTNRKRSFSLDIKSA----- 401
 414 AASLHGSSSEDALEYKPSSTLLDGN-----XLQCF-SP---VOELSEQTPETSPD 461
 402 -----YAPSRPDPGPGPPDPGPAKPLCKLSDPSGAALSLSPSPD-SFD 444

QY 462 KEASIPKLTQAPRPSDSQKRLHSVRTSSGTAQRLSLPRLSGSVEDNYHTSFLRG- 520
 DB 445 AAFEPAPRRPRPR-----PAGSPAR---SPAHSLG-----LNFGD 478
 QY 521 --LSTSOHLTKSAGLGLK-----WHSDLAPQSTPSLTSSWYATSSHPYSA 569
 DB 479 AARQTPRHGLSALAPGLPGPGAPAGCAWAPPLDSP--GTSPBDGCPWSP----- 529
 QY 570 SAIVGASASVAYSCSLPTCGDQVSVRRQRQSDSDRSRMHSPSPREKQPKRSCQ 629
 DB 530 GAQAGGVLFAFPFARAGAPGPGGSLDKRRBAARAEPRDRTKTPPEBPAPETQFKRRSCQ 589
 QY 630 MEFGESIMSEKSR-BELGKYGSQSPSGSMETILEVS 665
 DB 590 MEFGEG-MVEGRANGELIALGKQASPSGSGVEIVES 625

RESULT 9

Q7TS29

PRELIMINARY;

PRT; 665 AA.

AC Q7TS29
 DT 01-0CT-2003 (Tremblrel. 25, Created)
 DT 01-0CT-2003 (Tremblrel. 25, Last sequence update)
 DT 01-0CT-2003 (Tremblrel. 25, Last annotation update)
 DE Dusp8 protein.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strauberg R.L.; Reingold E.A.; Grouse L.H.; Derge J.G.;
 RA Klausner R.D.; Collins F.S.; Wagner L.; Shemen C.M.; Schuler G.D.;
 RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
 RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
 RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
 RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Schetz T.E.;
 RA Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
 RA Raha S.S.; Lounellano N.A.; Peters G.J.; Abramson R.D.; Mullaly S.J.;
 RA Bosak S.A.; McMan P.J.; McKernan K.J.; Matkovic J.A.; Gunaratne P.H.;
 RA Richards S.; Morley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulik S.W.;
 RA Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
 RA Fahey U.; Helton E.; Kettelman M.; Madan A.; Rodriguez S.; Sanchez A.;
 RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
 RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
 RA Rodriguez A.C.; Grimwood U.; Skalska U.; Myers R.M.; Buterfield Y.S.;
 RA Krzywinski M.I.; Skalska U.; Smallus D.E.; Schermer A.; Schein J.E.;
 RA Jones S.J.; Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strauberg R.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC052705; AAH52705.1; -
 DR SEQUENCE 665 AA; 69021 MW; 9166836A8835249F CRC64;

Query Match

Best Local Similarity 38.2%; Score 1304; DB 11; Length 665;
 Matches 319; Conservative 101; Mismatches 176; Indels 138; Gaps 26;

1 MAHEMIGTOIV-TERLVALLESGETEKVLLIDSRPFVEYNTSHLEAININCSKIMKRLQ 59
 1 MAGDRLPKRYMDAKKLASLRGPGGPIVYDSRFSVEYNSMHWLVSVMCCSKLVKRLQ 60
 60 QDKVITLIELIQAAXKHVDIDCSQKVVYVYDSSQVAVLSDCFITVLGLKLEKSFNSVH 119
 61 QOKVITLIELIQAAXKHVDIDCSQKVVYVYDSSQVAVLSDCFITVLGLKLEKSFNSVH 119

QY 120 LLAGFAEFCRPGCEGKSTLVPT-CISQPCLPVANIGTRILPMLYLGCORVANK 178
 DB 121 ILTGFAFESCFPGCEGKATLPMSLSQPCLPVSVGLTILPHLYLGSQKDVLMND 180
 QY 179 LIQONGIGYLNASYCPKPDFPESHFLRVPNVDSCEKILWLDKSDVFIRKAVASNG 238
 DB 181 LMTQNGISYVLANASNCPEKDFICESFMRIPINDVYCEKILPMLKSTIEFIDKATLSSC 240
 QY 239 CVLVHCLAGISRSATTAIAYIMKRMDSLDDEAYFVKEKPTISPNPNFGOLLDYEKKI 298
 DB 241 QVIVHCLAGISRSATTAIAYIMKRMDSLDDEAYFVKEKPTISPNPNFGOLLDYEKKI 300
 QY 299 KNOGTG--ASGPKSKLKLHLEKNEPEYPAVSEGGKSETPLSPPCADSAT--SEA- 350
 DB 301 KLAAAGTQDP-----HGTG-PFLMGPAGIPIPLRLP--PSTGESATSEEAATAAR 350
 QY 351 -----GORPHAVSVSVVQPSLLEDSPLVQALSGHLSADRLSDSNKLKSPSLD 403
 DB 351 EGSPSAGGDAPI-PSTAPATSAQ-----QGLRGHLSSDRLOQDTNRLKRSFSLD 399
 QY 404 IKSVYSASMAASLHGSSSEDALEYKSTLTIDGTNKLCOFSPVQELSEOTPEGPDK 463
 DB 400 IKS-AVAPSRPDPFGPPDGEAPKCLKDSFSGTGL-----PSPPDSP 445
 QY 464 EASIPK--KLQTAPEPDSQSKR--LHVSRTSSSGTAORSLPLHRSQSVEDNHTSFL 518
 DB 446 D-SVPECRRPPRRPPASPAASPAHGLGANGDTAKQ--TPRH----- 487
 QY 519 FGLSTSQOHLTKSAGL-----GLKGMHSDILAPQTSPLTSLTSMYFATESH---FY 567
 DB 488 -GLSA-----LSAPGLPQPGQPAQPGWVPLDSP--GTPSPDGMCFSPREGAQSGAVP 539
 QY 568 SA-----SAYTG-----SASYAVSCQLPTCGD 552
 DB 540 SAGRVASAGAPGPNSSSGGGGGGGGGGGGGSSSSSSSSSSSSSSSSSSSSSSSS 599
 QY 593 QVSVRRKQPSRADSRMSHEESPPEKQFKRRSCOMERGESIMSNRR-BELGVNG 651
 DB 600 SSDDLRR-----DRTWPEPEPAADQFKRRSCOMEREG-MWEGRAGBELALGK 651
 QY 652 QSSFGSGMEIIEVS 665
 DB 652 QTSFGSGVEIIEVS 665

RESULT 10
 Q9AG15 PRELIMINARY; PRT; 300 AA.

AC 09AG15;
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Map kinase phosphatase-M B2 isoform.
 GN DUSP16 OR 3830417M1R1X.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BA1B/c; Muslacharoen T., Johnson T.R., Kraft A.S.,
 RA Matsuguchi T., Yoshioka Y.,
 RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
 RT Activation in Macrophages.";
 RT Mol. Cell. Biol. 20:6999-7009 (2001).
 RL EMBL; AF345554; AK35055.1; -.
 DR HSSP; Q16828; MKP.
 DR MGD; MG1:1917936; Dusp16.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0005515; P:protein binding; IPI.
 DR GO; GO:0000188; P:inactivation of MAPK; IDA.

DR InterPro; IPR000340; D5_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS02066; RHODANES_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolase; Kinase.
 SQ SEQUENCE 300 AA; 33624 MW; 4C61846ACDF0F456 CRC64;

Query Match 30.4%; Score 1038; DB 11; Length 300;
 Best Local Similarity 66.5%; Pred. No. 2e-71;
 Matches 218; Conservative 13; Mismatches 23; Indels 74; Gaps 3;

QY 1 MAHEMTQITVERIVALLSEGEKTLIDSRPVEYNTSHILEAININCSKLMKRLQ 60
 DB 1 MAHEMTQITVESLVALLSEGEKTLIDSRPVEYNTSHILEAININCSKLMKRLQ 60
 QY 61 DKVLITELQHSARKKVIDIDCSQVVVYDQSDVASLSDDCFVLVLGLKGFNSVHL 120
 DB 61 DKVLITELQHSARKKVIDIDCSQVVVYDQSDVASLSDDCFVLVLGLKGFNSVHL 120
 QY 121 LLAGFAEFCRPGCEGKSTLVPTCISQPCLPVANIGTRILPMLYLGCORVANK 180
 DB 121 LA-----DLN 125
 QY 181 QONGIGYVNASYTCRKPDPFESHFLRVPNVDSCEKILPMLDKSDVFIRKAVASNGCY 240
 DB 126 QONGIGYVNASNTCKPDPFESHFLRVPNVDSCEKILPMLDKSDVFIRKAVASNGCY 185
 QY 241 LVHCLAGISRSATTAIAYIMKRMDSLDDEAYFVKE-----KPTISPNPNFG 289
 DB 186 LVHCLAGISRSATTAIAYIMKRMDSLDDEAYFVKE-----KPTISPNPNFG 245
 QY 290 QLLDYEKIKNOTG-----ASGPKS 309
 DB 246 AKWNLERALCRRTGPRSRWMAASPAS 273

RESULT 11
 Q8BZE4 PRELIMINARY; PRT; 206 AA.

AC 08BZE4;
 DT 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Truncated MAPK phosphatase 7 homolog.
 GN DEERTD213E.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
 RX MEDLINE=22354683; PubMed=1246651;
 RA The FANTOM Consortium.
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK035652; BAC29138.1; -.
 DR MGD; MG1:1098836; DEERTD213E.
 DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.

DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
SQ SEQUENCE 206 AA; 2256 MW; E7FDF5C01ADC2F73 CRC64;
Query Match 26.1%; Score 891; DB 11; Length 206;
Best Local Similarity 97.7%; Pred. No. 2.2e-60;
Matches 173; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAHEMIGTQVTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
DB 1 MAHEMIGTQVTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
QY 61 DKVITELIQHSAGHKVDIDCSQKVVVYDSSQDVASISDPCFLTVLLGKESFNSVHL 120
DB 61 DKVITELIQHSAGHKVDIDCSQKVVVYDSSQDVASISDPCFLTVLLGKESFNSVHL 120
QY 121 LAGGFAERSRCFPGICEGKSTLVPTCISQPCLPVANIPTRIILNLYLGCRDVLNK 177
DB 121 LAGGFAERSRCFPGICEGKSTLVPTCISQPCLPVANIPTRIILNLYLGCRDVLNK 177

RESULT 12
Q08ST19 PRELIMINARY; PRT; 657 AA.
ID Q08ST19
AC Q08ST19
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein F08B1.1a.
GN F08B1.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodetidae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Chisoe S.;
RT "The sequence of C. elegans cosmid F08B1.";
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR HSSP; Q16828; IMKP.
DR WormPep; F08B1.1a; CE27918.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR00387; TYR_phosphatase.
DR Pfam; PF00782; DSPc; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPc; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.

DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein; Hydrolase.
SQ SEQUENCE 657 AA; 71002 MW; A0D9153DE6326B43 CRC64;
Query Match 17.9%; Score 611; DB 5; Length 657;
Best Local Similarity 29.3%; Pred. No. 4.2e-38;
Matches 201; Conservative 104; Mismatches 198; Indels 184; Gaps 25;
QY 10 IYTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLQODPV----LI 65
DB 9 ISTCGLAALREAPDITLVVDCGFTERYNESHVRHSMNAFSLIRRLFENTLDNCL 68
QY 66 TELIQHSA-----KHKVDIDCSQKVVVYDQ-----SSQDVASL-----SSDCLTVLLGK 110
DB 69 HQMSSCGGCKTQKDEKLDL-----VLYAEDPKRGKRIASCNAPESFAKIMRVLRER 122
QY 111 LEKS--FNSYHLAAGFAERSRCFPGICEKS--TLVPTCISQPC--PVANIPTRIILP 164
DB 123 LEDTDKRSVMVLEGGKQFPAQYQPOLCESSSEGTRLPQSLQOPTGD--GITLITP 181
QY 165 NLVYGCORVYLNKELIQONGIGVYLNASYCPKPDFIPE--SHFLRVVNDSPCEKILPWL 223
DB 182 NLYIGSQIDSDTMDALDISVYINLSMTCPKSVCKEDKRMRIPVNDYQEKLSPTF 241

QY 224 DKAVDTEKAKASNGCVLVHCLAGISBATIAYIMKRMDSLDEAYRVYKRPITSP 283
DB 242 PMAYEFLEKCRRAKCLIHCLAGISRSPTAISYIMRYMKMGSDDAYRYVKERRPIS 301
QY 284 NFNPLQGLDYER-----KIKNOTGASGPKSKLKL--HLKRNPE----- 322
DB 302 NFNPMQGLLEYENVLIKDHVLDYNQASRPFRHNDYGPSPDLCPKPKASNSCVPFGST 361
QY 323 -----VPAVSEGGQKSEETPLSPPCADSAQAQRPVPAVSPVSPQSLLEDSP 375
DB 362 HDSSSPSPSPVSGSASSEPETSASASSSTNSA-----PSPMPTSR----- 405
QY 376 LVQALSGHLSDRLSDSNKLRKSFSLD-----KSVTSASMAASLHFFSSSEDALEY 429
DB 406 --QGTSGGVNVANG-----KRMMDLGLPHRPALGLPSRIIGTSV-----AEL 447
QY 430 YKPSITLDGTNKKCPSPVQGLSEQTP--ETSPDKEASIPKKLQIARPSDSQSKRLHS 486
DB 448 PSPSTELSRIS----FNGPPIALPSTPIINFTPCRNSTPIIV----- 486
QY 487 VRTSSSGTAQRSLSPLRSGSVEDNYHTSFRLGLSTSQOHLTKSAGLGLKGMHSDILAP 546
DB 487 --ASSSREVLITLTPPAASSS-----STSS----- 509
QY 547 QTSTPCLTSSWYATSSH-----FYSASAIYGGASYSYACSQLPTCGDQVYSVRR 600
DB 510 --SEPSFDFSPSSSSSIIVENPFPASTEVPASSSISTPSGSO----- 553
QY 601 QKPSDRADRSRSMHESPEKOFKRS 627
DB 554 STPASASSSAASCKRMKGFVKFSKKA 580

RESULT 13
Q06Q0S2 PRELIMINARY; PRT; 143 AA.
ID Q06Q0S2
AC Q06Q0S2
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Truncated MAPK phosphatase 7.
GN MKP7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Montpetit A., Boily G., Simmet D.;

RT "A detailed transcriptional map of the chromosome 12p12 tumor suppressor locus."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AV038927; AAK69770.1; -.
 DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR Pfam: PF00581; Rhodanese.1.
 DR PRINTS: PR01764; MARKPHPTASE.
 DR SMART: SM00450; RHOD.1.
 DR PROSITE: PS50206; RHODANES_3; 1.
 DR SEQUENCE 143 AA; 16077 MW; 5213A213AA7E5974 CRC64;
 SQ
 Query Match 17.8%; Score 610; DB 4; Length 143;
 Best Local Similarity 98.4%; Pred. No. 4.8e-39;
 Matches 144; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MAHEMIGTOIVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRRLQ 60
 DB 1 MAHEMIGTOIVTERLVALLSGTEKYLIDSRPFVEYNTSHLEAININCSKLMKRRLQ 60
 QY 61 DKVLTIELIQHSARKHVIDCSQKVVVYDQSSQDVASLSSDCEFLTVLLGKLEKSPNSVHL 120
 DB 61 DKVLTIELIQHSARKHVIDCSQKVVVYDQSSQDVASLSSDCEFLTVLLGKLEKSPNSVHL 120
 QY 121 LAGGFA 126
 DB 121 LAGADA 126
 RESULT 14
 Q8ST18 PRELIMINARY; PRT; 606 AA.
 AC Q8ST18;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Hypothetical protein F08B1.1b.
 GN F08B1.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Chisoe S.;
 RT "The sequence of C. elegans cosmid F08B1.";
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U31718; AAK68300.1; -.
 DR HSSP: Q16828; IMKP.
 DR WormRep: F08B1.1b; CE27919.
 DR GO: GO:0016787; P:hydrolase activity; IEA.
 DR GO: GO:0006138; P:protein tyrosine/serine/threonine phosphatase...; IEA.
 DR GO: GO:0006707; P:protein amino acid dephosphorylation; IEA.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSpC; 1.

DR SMART: SM00195; DSpC; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR KW Hypothetical protein; Hydrolase.
 DR SEQUENCE 606 AA; 64962 MW; 5B71795C6BD58058 CRC64;
 SQ
 Query Match 14.6%; Score 497.5; DB 5; Length 606;
 Best Local Similarity 28.9%; Pred. No. 2e-29;
 Matches 160; Conservative 77; Mismatches 159; Indels 157; Gaps 19;
 QY 123 GGFARFRCFPGLCRCKS--TLVPTCISQPCD--PVANIGPRILPNIYLGCQDVLNKE 178
 DB 86 GGFARFRCFPGLCRCKS--TLVPTCISQPCD--PVANIGPRILPNIYLGCQDVLNKE 178
 QY 179 LIQONGIGVYLAASYTCPEPDIPE-SHFLRPVNDSEFEKLLPWLKSDVDPEKAKASN 237
 DB 145 MLDALDISVIMLSMTCPEKVCIKEDKNFMRIPVNDSEFEKLLPWLKSDVDPEKAKASN 237
 QY 238 GCVLVHCLAGISRSATIAIAYIMKMDMLDAAYRVEKRPPTISNPNFLGQLDLYEK- 296
 DB 205 KKLCHLCLAGISRSPTLAIYIMRWKGSDDAIRYKERPSISNPNFLGQLDLYEK- 264
 QY 297 ---KIKNGTGASGPKSLKLL--HLEKNEP-----VPAVSEG 329
 DB 265 LKDHVLDVNGASRPHRMNDYGPDLCPKVPKASNSCVFPGTHDESSPSPSVSEG 324
 QY 330 GQKSTPLSPPCADATSEBAQRPVHPASVPSVSPVPSLLEDPLVALGHLHADR 389
 DB 325 SAASEPETSASAASSSTASA-----PSMESTSE-----QGTSGTVNVAG 366
 QY 390 LEDSNKLRFSGLDI-----KSVSYASAMASLHGFSSEDALEYKPSSTLLDGTNKC 443
 DB 367 ---KKNMTMDGLPHRPALGLPSRTGTSV-----AELPSSTELSRK--- 407
 QY 444 QSPVQELSEQTP---ETSPDKKEASIPKQLQTAPSPDSQSKRLHSVRTSSGTARSL 500
 DB 408 -FNGEALAPSPILNFTNPCFNSPIIV-----ASSREVLITLP 447
 QY 501 SPLHRSQSVEDNYHNSFLGLSTSOQHLTKSAGLGLKWHSDIAPQTSPLTSWYRA 560
 DB 448 TPAASSSS-----STS-----SEPSFDFSSPFS 470
 QY 561 TESSH-----FYSASAIYGSASAYSACQLPTCGDQVSVRRQKSDADSRSMH 614
 DB 471 SSSSIVENPFPASTEVPAGSSISTPESGQ-----STPASASSASAKC 516
 QY 615 EESPEKQPKRRS 627
 DB 517 RMKGFPKVPKRA 529
 RESULT 15
 Q7SZF3 PRELIMINARY; PRT; 367 AA.
 AC Q7SZF3;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Body;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

[illegible]

Search completed: June 21, 2004, 13:22:35
Job time : 53.2647 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 21:10:41; Search time 132.037 Seconds
(without alignments)
2794.987 Million cell updates/sec

Title: US-10-029-345A-109
Perfect score: 3418
Sequence: 1 MAHEMGTGTVTRVALLE.....LKNVSGSSRSGSMETIEVS 665

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cg2_1/USPTO.spool/US10029345/runcat 21062004 164146 29052/app_query.fasta_1.1294
-DB=Issued Patents NA -OEMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -INITs=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOR=6 -DELEXT=7

Database:

Issued Patents NA:
1: /cg2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cg2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cg2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cg2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cg2_6/ptodata/2/ina/PTCUS.COMB.seq:*
6: /cg2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3406	99.6	1998	4	US-09-816-494-3
2	3406	99.6	3544	4	US-09-816-494-1
3	1326	38.8	2377	4	US-09-920-668-3
4	480	14.0	2109	4	US-09-016-434-1135
5	480	14.0	2109	4	US-09-023-655-946
6	469	13.7	1830	4	US-09-557-921-1
7	453	13.3	1238	4	US-09-702-705-803
8	453	13.3	1238	4	US-09-736-457-803
9	453	13.3	1238	4	US-09-614-124B-803
10	453	13.3	1238	4	US-09-671-325-803
11	453	13.3	1238	4	US-09-589-184-803
12	451.5	13.2	2240	4	US-09-016-434-1100

13	450.5	13.2	2064	4	US-09-702-705-825	Sequence 825, App
14	450.5	13.2	2064	4	US-09-736-457-825	Sequence 825, App
15	450.5	13.2	2064	4	US-09-614-124B-825	Sequence 825, App
16	450.5	13.2	2064	4	US-09-671-325-825	Sequence 825, App
17	450.5	13.2	2064	4	US-09-589-184-825	Sequence 826, App
18	450.5	13.2	2109	4	US-09-702-705-826	Sequence 826, App
19	450.5	13.2	2109	4	US-09-736-457-826	Sequence 826, App
20	450.5	13.2	2109	4	US-09-614-124B-826	Sequence 826, App
21	450.5	13.2	2109	4	US-09-671-325-826	Sequence 826, App
22	450.5	13.2	2109	4	US-09-589-184-826	Sequence 826, App
23	447	13.1	1993	2	US-08-990-379-2	Sequence 2, Appl1
24	445	13.0	1238	2	US-08-530-290-11	Sequence 11, Appl
25	428.5	12.5	2000	4	US-09-016-434-1291	Sequence 1291, Ap
26	427.5	12.5	2303	4	US-09-922-146-3	Sequence 3, Appl1
27	425	12.4	944	4	US-09-371-671B-10	Sequence 10, Appl
28	405.5	11.9	1987	2	US-08-990-379-1	Sequence 1, Appl1
29	382.5	11.2	1619	4	US-09-702-705-801	Sequence 801, App
30	382.5	11.2	1619	4	US-09-736-457-801	Sequence 801, App
31	382.5	11.2	1619	4	US-09-614-124B-801	Sequence 801, App
32	382.5	11.2	1619	4	US-09-671-325-801	Sequence 801, App
33	382.5	11.2	1619	4	US-09-589-184-801	Sequence 801, App
34	382.5	11.2	1619	4	US-09-702-705-804	Sequence 804, App
35	382.5	11.2	1619	4	US-09-736-457-804	Sequence 804, App
36	382.5	11.2	1619	4	US-09-614-124B-804	Sequence 804, App
37	382.5	11.2	1619	4	US-09-671-325-804	Sequence 804, App
38	382.5	11.2	1619	4	US-09-589-184-804	Sequence 804, App
39	380	11.1	279	4	US-09-016-434-91	Sequence 91, Appl
40	295.5	8.6	539	4	US-09-389-681-311	Sequence 311, App
41	295.5	8.6	539	4	US-09-620-405B-311	Sequence 311, App
42	295.5	8.6	539	4	US-09-339-338-311	Sequence 311, App
43	295.5	8.6	539	4	US-09-433-826B-311	Sequence 311, App
44	295.5	8.6	539	4	US-09-604-287A-311	Sequence 311, App
45	295.5	8.6	539	4	US-09-834-759-311	Sequence 311, App

ALIGNMENTS

RESULT 1
US-09-816-494-3
; Sequence 3, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: (Meyers), Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494 - **ProtESID2**
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24 **ProtESID3**
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-816-494-3

Alignment Scores:
Pred. No.: 0
Score: 3406.00
Percent Similarity: 99.85%
Best Local Similarity: 99.70%
Query Match: 99.65%
DB: 4
Gaps: 0

US-10-029-345A-109 (1-665) x US-09-816-494-3 (1-1998)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
Db 1 ATGGCCCATGATGATGATGCAATTTGTACTGAGAGGTTGGCTGCTCGGAA 60
QY 21 SetG1Thrc1uyluvallleu1leAapSerArgProhbeValGluTyraAnThrSer 40

```
Db 61 ACTGGAACGAAAAGTCTGCTATTTAATACCGGCAATTTGGAAATCAATTCACATCC 120
Qy 41 HisIleuGluValAlaAsnIleAsnYserIleuMetLysArgArgLeuGln 60
Db 121 CACATTTTGAAGCATTAATATCAATGCTCAAGCTTATGAGCGAAAGTTGCAACAG 180
Qy 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 181 GACAAAGTTTAATATACAGAGCTCATCCAGCATTCAGCAACATTAAGGTGACATGAT 240
Qy 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 241 TGGAGTCAGAAAGTTGATGTTTACATCAAGCTCCAGATGTTGCTCTCTCTTCA 300
Qy 101 AspCysPheLeuThrValLeuLeuGluLysLysLeuGluLysSerPheAsnSerValHisLeu 120
Db 301 GACTGTTTCTCACGTAATCTCTGGGTAACTGGAGAAAGACTTCAACTCTGTTCACTG 360
Qy 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
Db 361 CTGCAAGGTGGGTTTGCTGAGTCTCTGTTGTTCCCTGGGCTCTGTAAGGAAATCC 420
Qy 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 421 ACTGTAAGTCCTACCTGCAATTTCTCAGCTTGCTTACCTGTTGGCAATTTGGGCAACC 480
Qy 161 ArgIleLeuProAsnLeuTyrLeuGluLysGlnArgAspValLeuAsnLysGluLeuIle 180
Db 481 CGAATTTCTCCCAATCTTATCTTGGCTGCCAGGAGATGTCCTCAACAGAGGTGATG 540
Qy 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
Db 541 CAGCAGATGGGATTTGGTTATGTTTAAATGCCAATACCTGTCCAACCTGACCTT 600
Qy 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
Db 601 ATCCCGAGTCTCAATTTCTCGTGCTGCTGATGATGACAGCTTTTGAAGAAATTTTG 660
Qy 221 ProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
Db 661 CCGGTGTGGAACAAATCGATGATTTCAATTGAAAGCAAAAGCTCCAAATGGATGTG 720
Qy 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db 721 CTAGTGCACTTTTATGCTGGGATCTCCGCTCCGCAACATGCTATGCTTACATCATG 780
Qy 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
Db 781 AAGAGATGAGCATGCTTTAGATGAACTTACAGATTTTGAAAGAAAAGAAAGCTTACT 840
Qy 281 IleSerProAsnPheAsnPheLeuGluLysLeuLeuAspLysGluLysLysIleLysAsn 300
Db 841 ATATCTCCAAACTTCAATTTCTGGGCCAACTCTGAGCTATGAGAAAGAAATTAAGAAC 900
Qy 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLysGluLysProAsn 320
Db 901 CAGACTGGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGAGAGAACCAAT 960
Qy 321 GluProValProAlaValSerGluGlyLysLysSerGluThrProLeuSerProPro 340
Db 961 GAACCTGCTCCTGCTGCTCAGAGGTTGACAGAAAGAGAGAGCGCCCTCAGTCCACC 1020
Qy 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyLysArgProValHisProAlaSerVal 360
Db 1021 TGTGCCCATCTGCTACCTCAGAGCGACAGACAAAGCGCCGTCGATCCCGCAAGGTG 1080
Qy 361 ProSerValProSerValGlnProSerLeuGluLysAspSerProLeuValGlnAlaLeu 380
Db 1081 CCCAGCGTGGCCAGCGTACAGCGCTGCTGTTAAGAGACAGCCGCTGCTACAGCGCTC 1140
Qy 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPhe 400
1141 AGTGGCTGACCTGTCCGAGACAGGCTGGAAGACAGATTAAGTCAAGCTTCTTC 1200
401 SerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
1201 TCTTGATATCAAAATGATTAATATTCAGCCAGCATGGCAGATCTTACATGGCTTC 1260
421 SerSerSerGlnAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAsn 440
1261 TCTCATCAGAAAGATGCTTTGGATTAATCAAACTTCCACTCTCTGATGGGACCAAC 1320
441 LysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
1321 AAGCTATGACAGTCTCCCTGTTACAGGACTATCGAGACAGATCCCGCAACCAAGTCT 1380
461 AspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
1381 GATTAAGAGAGAGCCACATCCCAAGAACTCAGACCCGAGGCTTTAGACAGCAG 1440
481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
1441 AGCAGAGATTCATTCGATCAGAACCCAGACAGACAGATGGCACCCGCAAGAGTCCCTT 1500
501 SerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheGly 520
1501 TCTCACTGCATCAGAGTGGAGCGTGGAGAGCAATTAACACACAGCTTCTTTTCGGC 1560
521 LeuSerThrSerGlnGlnHisLysLeuThrLysSerAlaGlyLeuGluLysGlyTyrPheHis 540
1561 CTTTCCACCGACGACGACCACTCAGAAATGCTGGCTGGGCTTAAAGGCTGGCAC 1620
541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
1621 TCGATATCTTGGCCCCCGACGACCTTACCCCTTCCCTGACGAGCTGTGATTTTGGC 1680
561 ThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSer 580
1681 ACAGAGTCTCACAATCTTCTGCTGCTCAGCCATCAACGAGGAGCTGCACTTACTCT 1740
581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnAlaTyrSerValArgArgArg 600
1741 GCTTACAGCTGACGACGACGCTGCCACTGCGGAGACCAATCTATCTGTCGCGAGCGG 1800
601 GlnLysProSerAspArgAlaAspSerArgSerTyrHisGluGluSerProPheGlu 620
1801 CAGAGCCAGATGACAGAGCTGACTCGCGCGAGCGGATGAAGAGACCCCTTTGAA 1860
621 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
1861 AAGCAGTTTAAAGCGAGAGCTCCAAATGGAATTTGGAGAGAGCATCATGTCAGAGAAC 1920
641 ArgSerArgGluGluLeuGluLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
1921 AGGTACGGGAAAGAGCTGGGAAAGTGGGAGTCACTTACTTTCGGGACAGATGAA 1980
661 IleIleGluValSer 665
1981 ATCATTTAGGCTTCC 1995
RESULT 2
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OR INVENTION: 3892 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR APPLICATION NUMBER: 2001-03-23
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589) ... (2583)
US-09-816-494-1

Alignment Scores:
Pred. No.: 0 Length: 3544
Score: 3406.00 Matches: 663
Percent Similarity: 99.85% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.65% Indels: 0
DB: 4 Gaps: 0

US-10-029-345a-109 (1-665) x US-09-816-494-1 (1-3544)
QY 1 MetAlaHieGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 589 ATGGCCCATAGAGATGATGGAACTCAAAATGTTACTGAGAGTGGCTGCTGCTGGA 648
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 649 AGTGGAAAGGAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 708
QY 41 HisIleGluGluAlaIleAsnIleAsnGlySerIleValMetIleValArgLeuGln 60
Db 709 CACATTTTGGAACTTAAATATCACTGCTCCCAAGCTTATGAAGCGAAGTTGCAAC 768
QY 61 AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Db 769 GACAAAGTGTAAATATACAGAGCTATCAGCAATTCAGCAACCAAGTGAACATGAT 828
QY 81 CysSerGlnIleValValIleValIleValIleValIleValIleValIleValIle 100
Db 829 TGCAGTCAGAGAGTGTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 888
QY 101 AspCysPheLeuThrValIleLeuGlyValLeuGlyValLeuGlyValLeuGlyVal 120
Db 889 GACTGTTTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 948
QY 121 LeuAlaGlyValIleAsnIleAsnGlySerIleValMetIleValArgLeuGln 140
Db 949 CTTCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyPro 160
Db 1009 ACTCTAGTCCCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
QY 161 ArgIleLeuProAsnLeuIleGlyValLeuGlyValLeuGlyValLeuGlyValLeu 180
Db 1069 CGAATTCCTCCCACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1128
QY 181 GlnGlnAsnGlyIleGlyValLeuAsnIleAsnGlyValLeuGlyValLeuGlyVal 200
Db 1129 CACACAGATGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1188
QY 201 IleProGlnSerIlePheLeuArgValProValAsnAspPheCysGlyValIleLeu 220
Db 1189 ATCCCGCAGTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
QY 221 ProTrpLeuAspIleSerValAspPheIleGluValAlaValAsnGlyValIleVal 240
Db 1249 CCGGTGTGGACAAATCAGTATGATTTCTATGAGAAACAAAGCTCCATGATGTGTT 1308
QY 241 LeuValHisGlySerLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIle 260
Db 1309 CTAGTGCACGTGTAGTGGAGATCTCCGCTCCGCCACCAATCGCTATCGCTATCAT 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaIleValArgPheValIleGlyValArgPro 280
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 1369 AAGAGATGACATGCTTTAGATGAGCTTACGATTTGTGTAAAGAAAAAGACCTACT 1428
QY 281 ILeSerProAsnPheAsnIleLeuGlyIleLeuLeuAspTrpGluValIleValAsn 300
Db 1429 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGACTATGAGAAAGATTAAAGAAC 1488
QY 301 GlnThrGlyAlaSerGlyProIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 320
Db 1489 CAGACTGAGACATCAGGCGCCAAAGACCAACTAACCTGCTGCACTGAGAAAGCCAAAT 1548
QY 321 GluProValProAlaValSerGluGlyIleGlyValLeuSerGluThrProLeuSerPro 340
Db 1549 GAACCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1608
QY 341 CysAlaAspSerAlaThrSerGluAlaIleGlyValArgProValHisProAlaSerVal 360
Db 1609 TGTGCGACTGTGCTACCTCAGAGGAGGAGCAAGAGCCGCTGATCCCGCAGCGTG 1668
QY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
Db 1669 CCGAGCTGCCAGCTGCGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1728
QY 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIleValArgSerPhe 400
Db 1729 AGTGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1788
QY 401 SerLeuAspIleIleValSerValSerTrpSerAlaSerMetAlaIleSerLeuHis 420
Db 1789 TCTCTGATATCAATCAAGTTCATATTCAGCCAGATGGACATCTTACATGCTTTC 1848
QY 421 SerSerSerGluAspAlaLeuGluIleValIleValIleValIleValIleValIle 440
Db 1849 TCTCTATCAGAGATGCTTGGAAATCTAACAACTTCCACTGCTGATGGAGCAAC 1508
QY 441 LysLeuCysGlnPheSerProValGlnIleLeuSerGluGlnThrProGluThrSerPro 460
Db 1909 AAGCTATGCAAGTTCCTCCCTGTTAGAACTATCGAAGAGACTCCCGCAACAGTCT 1968
QY 461 AspIleGluGluAlaSerIleProIleValLeuGlnThrAlaArgProSerAspSerGln 480
Db 1969 GATTAAGAGAGAACCCAGATCCCAAGAGCTGCAACCGCGCTTCAGACAGCCAG 2028
QY 481 SerIleArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
Db 2029 AGCAGCGATGCTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2088
QY 501 SerProLeuHisArgSerGlySerValGluAspAsnTrpHisThrSerPheLeuPheGly 520
Db 2089 TCTCCACTGATGAGATGGAGCGTGAAGCAATTCACACAGCTTCTTTTCGGC 2148
QY 521 LeuSerThrSerGlnGlnHisLeuThrIleSerAlaGlyLeuGlyValLeuGlyVal 540
Db 2149 CTTTCCACAGCCAGCAGACCTCAGAAATGCTGCGCTGCGCTTAAAGGCTGGCAC 2208
QY 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpIlePheAla 560
Db 2209 TCGGATATCTGGCCCCCAGACCTTACCTTCCCTGACACAGAGCTGATTTTGGC 2268
QY 561 ThrGlnSerSerHisPheTrpSerAlaIleIleValIleGlyValSerAlaSerTrpSer 580
Db 2269 ACAGAGTCCCTCAGACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2328
QY 581 AlaTrpSerCysSerGlnLeuProThrCysGlyValArgGlnValIleValIleVal 600
Db 2329 GCTTACAGCTGACGAGCTGCCACTTCCGAGAACCAAGTCTATCTGCTGCTGCTG 2388
QY 601 GlnIleProSerAspArgAlaAspSerArgArgSerTrpHisGluGluLeuSerProPheGlu 620
Db 2389 CAGAGCCAGATGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2448
QY 621 LysGlnPheIleValArgArgSerCysGlnMetGluPheGlyIleValSerIleMetSerGluAsn 640
Db 2449 AAGCAGTTTAAAGCAGAGCTGCCAAATGGAATTTGAGAGAGCATGTCATGTCAGAAAC 2508
```


DB 1536 CCCGCGACAGCTCGGC-----CTGACCTCGGCGAT 1568
QY 521 -----LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyGly 538
DB 1569 GCGCGCCGCGACCTCGCGGCAAGCGCTCTCGCCCTCGCGCCGCGCGCGCGCGC 1628
QY 539 -----TrpHisSerAspIleLeuAlaProIleThrSer 549
DB 1629 CTTGGCCAGCGCGCGCGCGCGCGCGCTGGGACCGCGCTTACCTCCCA-----GGC 1682
QY 550 ThrProSerLeuThrSerSerTrpTyrrPheAlaThrGluSerSerHisPheTyrrSerAla 569
DB 1683 AGCCCGTCGCGCGACGCGCGCTGTGTCTTCAAGCCCGAG----- 1721
QY 570 SerAlaIleTyrrGlySerAlaSerTyrrSerAlaTyrrSerGlySerGlnLeuProThr 589
DB 1722 GCGCGACAGCGCGCGCGCGCGGTGCTGTTCGCGCCCTTCGCGCGCGCGCGCGCGGGA 1781
QY 590 CysGlyAspGlnValTyrrSerValArgArgGlnLysProSerAspArgAlaAspSer 609
DB 1782 CCAAGCGCGCGCGACGACCTCGCGCGCGGAGGACGAGCGCTGAGCCCGGAGCGG 1841
QY 610 ArgArgSerTrpHisGlnGluSerProPheGlnLysGlnPheLysArgArgSerCysGln 629
DB 1842 CGGACCGCGCTGCGCGCGCGCGCGCGCGGAGCGAGTTCAAGCGCGCGCACTGCGCA 1901
QY 630 MetGluPheGlyGluSerIleMetSerGluAsnArgSerArg---GluGluLeuGlyLys 648
DB 1902 ATGAGAGTTCAAGAGGAGCGC---ATGATGAGAGGCGCGCGCGCGCGCGCGCGCGCC 1958
QY 649 ValGlySerGlnSerPheSerGlySerMetGluIleIleGluValSer 665
DB 1959 CTGGGCAAGAGCGAGCTTCTCGGCGAGCGTGAAGTCAATCAAGGTGTCTC 2009

RESULT 4

US-09-016-434-1135
Sequence 1135 Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Sellhammer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016.434
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1135:
SEQUENCE CHARACTERISTICS:

LENGTH: 2109 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 91418933
US-09-016-434-1135

Alignment Scores:
Pred. No.: 8,466-40 Length: 2109
Score: 480.00 Matches: 121
Percent Similarity: 51.62% Conservative: 70
Best Local Similarity: 32.70% Mismatches: 126
Query Match: 14.04% Indels: 53
Gaps: 9

US-10-029-345A-109 (1-665) x US-09-016-434-1135 (1-2109)

QY 19 LeuGluSerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrrAsn 38
DB 430 CTGAGCTGCGCAACGAGCGCGCTGTGTGATGATGATGCGCGCGCGAGACTATACGAG 489
QY 39 ThrSerHisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeu 58
DB 490 TCGTCGACATGACGATCGCGCATCAACGCGCATCCCGGATCATGCTCGCGCGCTG 549
QY 59 GlnGlnAspLysValLeuIleThrGluLeuIleGlnHisSerAla---LysHisLysVal 77
DB 550 CAGAAAGGTAACTACCTGCGCGCGCGCTTTCACCGCGCGAGGACCGGACCGCTTC 609
QY 78 AspIleAspCys---SerGlnLysValValValTyrrAspGlnSerSerGlnAsp---Val 95
DB 610 ACCCGCGCTGTGAGCAGCAGACAGATGATGCTTACGAGAGAGAGCAGCGAGCTGAGAC 669
QY 96 AlaSerLeuSerSerAspCysPheLeuThrValLeuGlyLysLeuGluLysSerPhe 115
DB 670 GAGAAATACGCGCGCGCGAGCTGCTGCGCGCTGCTCTCAAGAGCTCAAGACGAGGCG 729
QY 116 AsnSerValHisLeuLeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeu 135
DB 730 TCGCGCGCTTCACTGAGAGGTGCTTCAGAACTTCAAGCCGAGTTCTCCCTGAT 789
QY 136 CysGluGlyLysSerThrLeuValProThrCys---IleSerGlnProCysLeuProVal 154
DB 790 TGGAG-----ACCAATTAAGCGGCTGTGAGCAGAGCTCGCGCGCTTGCAAG 843
QY 155 AlaAsnIleGly----- 158
DB 844 CTGGGCTCGGGGCGCTGCGGATCACTGACTCTTCTCGGACATCGAGTGAAGCTT 903
QY 158 ----- 158
DB 904 GACCGAGACCCCAATGATGCAACAGACTCGGATGATCCGCTGCAACAGCAGAGCT 963
QY 159 -----ProThrArgIleLeuProAsnLeuTyrrLeuGlyCysGlnArgAspValLeuAsn 176
DB 964 TCGTTCACAGTGAAGATCTTCCCTTCTCTCTGCGGTGCGCAAGACTCCACCAAC 1023
QY 177 LysGluLeuIleGlnAsnGlyIleGlyTyrrValLeuAsnAlaSerTyrrThrCysPro 196
DB 1024 TTGAGCTGTGGAGAAATTCGCGATCAAGTACATTTGAAGTCACCCCAATTTGGCG 1083
QY 197 LysPro---AspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPhe 215
DB 1084 AATCTCTTGAGAACGAGAGGTTTAATTAACAAACAAATCCCATCTCGATCACTG 1143
QY 216 CysGluLysIleLeuProThrLeuAspLysSerValAspPheIleGluLysAlaLysAla 235
DB 1144 AGCCAAACCTGTCAGTTTTCCTGAGGCGCATTTCTTCAATGATGAAGCCGCGGC 1203
QY 236 SerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAla 255

Db 1204 AAGAACTGTGGTGTCTGTGACATATTCCTGGCTGGACCGCTCATGTCAGTCTGAGT 1263
QY 1256 ILeaATyrlleMellysArMecAspMetSerleuaspbluAlaTyraGpHeValys 275
Db 1264 GTGGCTTACCTTAATGCAAGAGCTAATCTGTGATGAAACATGCTTAATGACATTTGTCAAA 1323
QY 276 GlulysArpProthrIleSerProAsnPheAsnPheleuylGlnleuAspTyrglu 295
Db 1324 ATGAAAAAATCCACATATCCCTTAATCTCACTTCATGGGTGAGCTGGAGCTTCGAG 1383
QY 296 LysLysIleLysAsnGlnThrGlyAlaSerGlyPro-----LysSerLysLeuLys 312
Db 1384 AGGACCGTGG-----GAGCTCAGCAGCCCATGTGACAAACAGGGTTCAGCAGCAG 1431
QY 313 LeuLeuHisLeuGluLysProAsnGluProValProAlaValSerGluGlyGlnLys 332
Db 1432 CAGCTGATTTTACCAACCCCTTCCAAACGAAATGATACAGGTGAGTCTCTGCAATCT 1491
QY 333 Ser-GluThrProLeuSerProProCys 341
Db 1492 ACGTGAAGAGCCCAACACCCCTCTCTGTC 1519
RESULT 5
US-09-023-655-946
Sequence 946, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 946:
SEQUENCE CHARACTERISTICS:
LENGTH: 2109 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1418933
US-09-023-655-946
Alignment Scores: 8.46e-40 Length: 2109
Pred. No.: 1

Score: 480.00 Matches: 121
Percent Similarity: 51.62% Conservative: 70
Best Local Similarity: 32.70% Mismatches: 126
Query Match: 14.04% Indels: 53
DB: 4 Gaps: 9
US-10-029-345A-109 (1-665) x US-09-023-655-946 (1-2109)
QY 19 LeuGluSerGlyThrGluLysValLeuLeuIleAspSerArpProPheValGluTyraAsn 38
Db 430 CTGGAAGGTAACTGCGCGCGCGCGCGCTTTCACCGCGCGAGACCGGACCGCTTC 489
QY 39 ThrSerHisLeuGluAlaIleAsnIleAsnCysSerLysLeuMellysArpGluLeu 58
Db 490 TGTGGCACAATCGATCGGCGCATCAACGTGGCCATCCCGGCAATCATGCTGCGCGCTCG 549
QY 59 GlnGlnAspLysValLeuIleThrGluLeuIleGlnHiserAla---LysHisLysVal 77
Db 550 CAGAAAGGTAACTGCGCGCGCGCGCGCTTTCACCGCGCGAGACCGGACCGCTTC 609
QY 78 AspIleAspCys---SerGlnLysValValValTyraSpLinsSerGlnAsp---Val 95
Db 610 ACCCGCGCTGTGCGACCGACACACAGTGTGCTTACGACGAGGACGACGACCTGGAAC 669
QY 96 AlaSerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPhe 115
Db 670 GAGAAATCGGCGCGCGCGCGATGCTGTGCGCGCTGCTCAAGAAAGCTCAAGACGAGGC 729
QY 116 AsnSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerArpCysPheProGlyLeu 135
Db 730 TGCAGGCGCTTCTACCTGAGAAAGGTGCTTCAGTAAGTTCAGACCGAAGTTCCTCGCAT 789
QY 136 CysGluGlyLysSerThrLeuValProThrCys---IleSerGlnProCysLeuProVal 154
Db 790 TGGAGC-----ACCATTAAGACGCTGTGTACAGACAGCTCCCGCGTTGGCAGTG 843
QY 155 AlaAsnIleGly----- 158
Db 844 CTGGGCGCTCGGGGCGCTGGGATCAGCTCTGACTTCTTCTCGGACATGAGTTCAGCTT 903
QY 158 ----- 158
Db 904 GACCGAGACCCCAATAGTGCACAGACTCGAGTATGCCGTGCCAAGCAGCGCT 963
QY 159 -----ProThrArgIleLeuProAsnLeuTyrlleuGlyCysGlnArgAspValLeuAsn 176
Db 964 TCCCTCCCGAGTGAATCTTCCCTTCTTAATGAGTGGCTGGCCAAAGACTCCACCAAC 1023
QY 177 LysGluLeuIleGlnGlnAsnGlyIleGlyTyraValLeuAsnAlaSerTy-ThrCysPro 196
Db 1024 TTGAGACGTGTGAGGAATCGGCAATCAATCTTGAAAGTCAAGCCCCCAATTTGCCG 1083
QY 197 LysPro---AspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPhe 215
Db 1084 AATCTCTTGAACAGCAGAGAGTTTAAATACAGCAAAATCCCATCTCGATCACTGG 1143
QY 216 CysGluLysIleLeuProThrLeuAspLysSerValAspPheIleGluLysAlaValAla 235
Db 1144 AGCAGAAACCTGTCCCACTTTTCCCTGAGGCCAATTTCTTCAATAGTAAAGCCGGGCGC 1203
QY 236 SerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAla 255
Db 1204 AAGAACTGTGGTGTCTGTGACATATTCCTGGCTGGACCGCTCAGTCACTGAGTACT 1263
QY 256 ILeaATyrlleMellysArMecAspMetSerleuaspbluAlaTyraGpHeValys 275
Db 1264 GTGGCTTACCTTAATGCAAGAGCTAATCTGTGATGAAACATGCTTAATGACATTTGTCAAA 1323
QY 276 GlulysArpProthrIleSerProAsnPheAsnPheleuylGlnleuAspTyrglu 295
Db 1324 ATGAAAAAATCCACATATCCCTTAATCTCACTTCATGGGTGAGTCTCTGCAATCTGAG 1383
QY 296 LysLysIleLysAsnGlnThrGlyAlaSerGlyPro-----LysSerLysLeuLys 312

```
Db 1384 AGAGAGCTG-----GAGCTGAGAGCCCATGTGACAAACAGGGTTCAGCAG 1431
Qy 313 LeuLeuH1sLeuGluYsProAngluProValProAlaValSerGluGlyGluLys 332
Db 1432 CAGCTGATTTTACACACCCCTTCCACCAAGATGTATACAGGTGAGCTCTGTGCATCT 1491
Qy 333 Ser-GluThrProLeuSerProProCys 341
Db 1492 ACGTAAAGACCCACACCCCTCTGTC 1519

RESULT 6
US-09-557-921-1
; Sequence 1, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.416
; CURRENT FILING DATE: US/09/557,921
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-557-921-1

Alignment Scores:
Pred. No.: 9,28e-39 Length: 1830
Score: 469.00 Matches: 107
Percent Similarity: 57.72% Conservative: 65
Best Local Similarity: 35.91% Mismatches: 96
Query Match: 13.72% Indels: 30
Gaps: 7

US-10-029-345A-109 (1-665) x US-09-557-921-1 (1-1830)
Qy 27 LeuLeuH1sLeuSerArgProPheValGluThrAsnThrSerHisLeuGluAlaIle 46
Db 603 GTATCATTTGACGAGCGCCCTTATGAGTACAAAGATCATTCAGAGAGCTGTC 662
Qy 47 Asn1LeuAncYsSer---LysLeuMetLysArgLysGluGlnGlnAspLysValLeuIle 65
Db 663 CACATTAACTGTGCGCATGATCAGCGCGGAGACTGCGAGCGGCAAGATCATCTGTC 722
Qy 66 ThrGluLeuIleGln---HisSerAlaLysHisLysValAspLeuAspCysSerGlnLys 84
Db 723 CTAGACTTGATTTCTCTAGAGGAAGCAAGCATCTTTCAGAGAGATCTTTCCAAAGAA 782
Qy 85 ValValVal1YrAspGlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeu 104
Db 783 ATTATATTTTATGTATGAGTATACCAATGACCAAGCGAGTGAATCCCTCCAGCCACTT 842
Qy 105 ThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeuLeuAlaGlyLys 124
Db 843 CACATTAAGCTCTGAGCTCTGAGAGAGAAAGCAAAAGCACTTGAGTTGAAGGTGGA 902
Qy 125 PheAlaLupPheSerArgCysPheProGlyLeuGluGluGlyLys-----139
Db 903 CTTAGTGTGTTTAAACAAACATGAAACCTGTGACCAATCTCTCCAGCTCCAAAGAG 962
Qy 140 -----SerThrLeuValProThrCysIleSer 148
Db 963 TGCCGGAGAGTGGGGGGGGCGCATCGCGGCTCGAGCTTGCTGACTCCAGCCCATC--- 1019
Qy 149 GlnProCysLeuPro---ValAlaAsn1IleGlyProThrArgLysLeuProAsnLeuTyr 167
Db 1020 ---CCCAACCAACCCCTGACATCGAAGACGTGAGCTGACCCCAATCTTGCTCTGTC 1076
Qy 168 LeuGlyCysGlnArgAspValLeuAsnLysGluLeuIleGlnGlnAsnGlyLysIleTyr 187
```

```
Db 1077 CTTCGCAATGAGAGGATGTCTGACAGACCTGAGACACATGAGCGGCTGAACATCGGCTAC 1136
Qy 188 ValLeuAsnAlaSer-----TyrThrCysProLysProAspPheIlePro 202
Db 1137 GTCATCAAGTCAACCACTCATCTTCCCTTACCATATGAGAAAGCCCTGTC----- 1190
Qy 203 GluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeuProTyr 222
Db 1191 -----ACTACAGAGCGGCTGCGCAGCACTGACAGCAACAGCAAACTGCGGCACTAC 1244
Qy 223 LeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysValLeuVal 242
Db 1245 TTGGAAGAGCGCTTTGATGTTCAATGAGAAAGTACACAGTGGGAAGGCGCTTCATC 1304
Qy 243 HisCysLeuAlaGlyLysSerArgSerAlaThrIleAlaIleAlaTyrIleMetLysArg 262
Db 1305 CACTGCAGGCTGGGGGTGTCCTCCGCGCACCATGCTGCTTACTTGATGAAGCAC 1364
Qy 263 MetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThrIleSer 282
Db 1365 ACTGGATGACCACTGACTGATGTATTAATTGTCAAGGCAAGCAACCAATTATCTCC 1424
Qy 283 ProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
Db 1425 CCAACTTAATCTTCAATGGGCGCATGCTGAGAGTTCGAGGAAGACCTAAACAAAC 1478

RESULT 7
US-09-702-705-803
; Sequence 803, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Derrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT FILING DATE: US/09/702,705
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 803
; LENGTH: 1238
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-803

Alignment Scores:
Pred. No.: 2.14e-37 Length: 1238
Score: 453.00 Matches: 120
Percent Similarity: 50.52% Conservative: 73
Best Local Similarity: 31.41% Mismatches: 143
Query Match: 13.25% Indels: 46
Gaps: 13

US-10-029-345A-109 (1-665) x US-09-702-705-803 (1-1238)
Qy 19 LeuGluSerGlyThrGluLysValLeuLeuAspSerArgProPheValGluTyrAsn 38
Db 168 CTCGCGAGGGGC---GCAAGTGCCTGCTGCTGGACTGAGACCGTTCCGCGGCGACAGC 224
Qy 39 ThrSerHis1IleLeuGluAlaIleAsn1LeuAncYsSerLysLeuMetLysArgAlaGlu 58
Db 225 GCGGGCTAATCTCTAGGTGTGTCACAGTGGCGCTGTAACCACTTCGCGGCGGCGG--- 281
Qy 59 GlnGlnAspLysValLeuIleThrGluLeuIle-----GlnHisSerAlaLysHisLys 76
```

```

Db      282 GCTAAGGCTCCGTGAGCGAGCATCTCTCCCGGAGAGAGATACGGCCCGC 341
Qy      77 ValAspIleAspCysSerGlnIlyValValValTyraAspGlnSerGlnAspValAla 96
Db      342 TTGCGCTCCGGCTTACTGCGCGGTCTGCTCTACGACGCGCAGCCGCGCGAG 401
Qy      97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyIlySleuGlnIlySerPhe--- 115
Db      402 AGCTTCGCGAGGAGCAGCAGCGTGCCTGGTGGGAGGCGCGCGCGCGAG 461
Qy      116 ---AsnSerValHisLeuLeuAlaGlyIlyPheAlaGlnPheSerArgCysPheProGly 134
Db      462 CGCACCAGCATCTGCTCTCAAAAGCGGTATAGAGGTTTCTCTCGAGTACCCAGAA 521
Qy      135 LeuCysGlnIlyIlySerThrIleu-----ValProThrCysIleSer 148
Db      522 TTCTGTTCTAAACCAAGCCCTGGACCATCCACCCTCCCGGTTCCCGCAGCGCACA 581
Qy      149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db      582 GAGCCCTTGAACCTGAGACTGCTCTGTGGAGACCCCACTACAGACAGCAGAGGCTCT 641
Qy      160 ThrArgIleLeuProAsnLeuTyLeuGlyCysGlnArgAspValLeuAsnIlySleu 179
Db      642 GTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 701
Qy      180 IleGlnIleAsnIlyIleGlyTyValLeuAsnAlaSerTyThrCysProIlyProAsp 199
Db      702 CTGAGACCCCTGGGACATACGCGCTGTGTGAATGTCTCTCGAGACGCCCAAC--CAC 758
Qy      200 PheIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIlySle 219
Db      759 TTTCAGAGACATATCATCAAGTCACTCCGATCGAAGAAATACACAGGCGCACATC 818
Qy      220 LeuProTyLeuAspIlySerValAspPheIleGlnIlyAlaIlyAlaSerAsnIlyCys 239
Db      819 AGCTCTCGGTTTCATGAGCCATAGAGTATCATGATCCGTAAAGACATCCCGTGGCGC 878
Qy      240 ValLeuValHisCysLeuAlaGlyIlyIleSerArgSerAlaThrIleAlaIleAlaTyrlle 259
Db      879 GTCTGTGTGACATCGCAGCGCGCATCTGCGGTGCGGCGCACCATCTGCGTGGCTACCTG 938
Qy      260 MetIlyArgMetAspMetSerLeuAspGlnAlaIlyTyraGpPheValIlyGlnIlyArgPro 279
Db      939 ATGATGAAGAAACGGGTGAGGCTGAGAGAGGCTTCGAGTTCTGTAAACAGCGCGCAGC 998
Qy      280 ThrIleSerProAsnPheAsnPheLeuGlyIlyLeuLeuAspTyTyGlnIlySleIlyS 299
Db      999 ATCATCTCGCCCAACTTCACTTCAATGGGCGAGCTGCGAGTTCCAGAGTCCAGAGTCTG 1058
Qy      300 AsnGlnThrGlyAlaSerGlyProIlySerIlySleuIlySleuGlnIlySleuGlnIlyPro 319
Db      1059 GCCACGCTCTGTGTCGCGAGGCTGCTAGC-----CCC 1091
Qy      320 AsnGlnProValProAlaValSerGlnIlyGlnIlySerGlnIlyProIlyProIlySerPro 339
Db      1092 TCGGAGCCCTG-----GGGAGCGCGGCAAGACCCCGCCAGCCCGC 1133
Qy      340 ProCysAlaAspSerAlaThrSerGlnAlaAlaGlyIlyGlnArgProValHisProAlaSer 359
Db      1134 -----ACCTTCGACGTGCTCTTCAAGCTTTCGCGC--TCCGTGGCGC 1172
Qy      360 ValProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAla 379
Db      1173 GTGAGCTCGCGCCCAAGCAGCGCTCTACCTG-----CACAGCCCATCACCACTCT 1226
Qy      380 LeuSer 381
Db      1227 CCCAGC 1232

```

RESULT 8
US-09-736-457-803

```

/ Sequence 803, Application US/09736457
/ Patent No. 6509448
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tonglong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedrick, Tom
/ APPLICANT: Carter, Patrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mammion, Jane
/ APPLICANT: Fan, Liqun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C15
/ CURRENT APPLICATION NUMBER: US/09/736,457
/ NUMBER OF SEQ ID NOS: 1864
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 803
/ LENGTH: 1238
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ US-09-736-457-803

```

Alignment Scores:

```

Pred. No.: 2.14e-37 Length: 1238
Score: 453.00 Matches: 120
Percent Similarity: 50.52% Conservative: 73
Best Local Similarity: 31.41% Mismatches: 143
Query Match: 13.25% Indels: 46
Gaps: 13

```

US-10-029-345A-109 (1-665) x US-09-736-457-803 (1-1238)

```

Qy      19 LeuGlnSerGlyThrGlnIlyValLeuLeuIleAspSerArgProPheValGlnTyraSn 38
Db      168 CTGCGAGAGCGC---GCGAAGTCTCTGCTGAGACTGCAACAGCGTCTCGGCGCAGC 224
Qy      39 ThrSerHisIleLeuGlnAlaIleAsnIleAsnCysSerIlySleuMetIlyArgIlySleu 58
Db      225 GCGGCTCATCTCTGATGTTGCTGTCAGCGTCAACGCGCTGAACACATCTGCGCGCGG--- 281
Qy      59 GlnGlnAspIlyValLeuIleThrGlnIleIle-----GlnHisSerAlaIlyHisIlyS 76
Db      282 GCTAAGGCTCCGAGAGCTGAGACCAAGTCTGCCCGCGAGGAGAGTACCGCGCGC 341
Qy      77 ValAspIleAspCysSerGlnIlyValValValTyraAspGlnSerGlnAspValAla 96
Db      342 TTGCGCTCCGGCTTACTGCGCGGTCTGCTCTACGACGCGCAGCCGCGCGAG 401
Qy      97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyIlySleuGlnIlySerPhe--- 115
Db      402 AGCTTCGCGAGGAGCAGCAGCGTGCCTGGTGGGAGGCGCGCGCGCGAG 461
Qy      116 ---AsnSerValHisLeuLeuAlaGlyIlyPheAlaGlnPheSerArgCysPheProGly 134
Db      462 CGCACCAGCATCTGCTCTCAAAAGCGGTATAGAGGTTTCTCTCGAGTACCCAGAA 521
Qy      135 LeuCysGlnIlyIlySerThrIleu-----ValProThrCysIleSer 148
Db      522 TTCTGTTCTAAACCAAGCCCTGGACCATCCACCCTCCCGGTTCCCGCAGCGCACA 581
Qy      149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db      582 GAGCCCTTGAACCTGAGACTGCTCTGTGGAGACCCCACTACAGACAGCAGAGGCTCT 641
Qy      160 ThrArgIleLeuProAsnLeuTyLeuGlyCysGlnArgAspValLeuAsnIlySleu 179
Db      642 GTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 701
Qy      180 IleGlnIleAsnIlyIleGlyTyValLeuAsnAlaSerTyThrCysProIlyProAsp 199

```

```

Db      702 CTGAGACCGCCGGGACATCAAGGCTGTGATGTCCTCCGAGACTGCCAAAC---CAC 758
Qy      200 PheIIeProgluSerHisPheLeuArgValProValAsnAspSerPheCysGluValIle 219
Db      759 TTGGAAGACACTATCACTGACAGTGCATCCAGTGAAGATTAACCAAGCGCCGACATC 818
Qy      220 LeuProTfPheAspIleAspSerValAspPheIIeGluValAlaValAsnAsnGlyCys 239
Db      819 AGCTCTGTTGATGAGAGCCATAGAGTACATGATCCCTGAAGAGACTGCCGCGGCGC 878
Qy      240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIle 259
Db      879 GTGTGTGTGACACTCCAGGCGGCGCATCTCGCGGTCCGACCAACATCTGCTCCGCTACCTG 938
Qy      260 MetLeuArgMetAspMetSerLeuAspGluAlaIleArgPheValIleGluValArgPro 279
Db      939 ATGATGAAGAACCGGAGCTGAGGCTGAGAGAGGCTTCGAGTTCGTTAAGCAGCGCCGAGC 998
Qy      280 ThrIleSerProAsnPheAsnPheLeuGlyGluLeuAspIleArgIleValIleVal 299
Db      999 ATCATCTCGCCCAACTTCAAGCTTCATGGGAGCTGCTGAGTTCGAGTCCCAAGTGTG 1058
Qy      300 AsnGlnThrGlyAlaSerGlyProIleSerIleValLeuValIleGluValPro 319
Db      1059 GCCACGCTCTGTGCTGCGGAGGCTGCTAGC-----CCC 1091
Qy      320 AsnGluProValProAlaValSerGluGlyGluIleValSerGluThrProLeuSerPro 339
Db      1092 TCGGAGACCCCTG-----GGGAGAGCGGGCAAGACCCCGCGCACCC 1133
Qy      340 ProCysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSer 359
Db      1134 -----ACCTCGCACTGCTGCTTTCAGCTTTCGCGTC---TCGTTGGGC 1172
Qy      360 ValProSerValProSerValGlnProSerIleuGluAspSerProLeuAlaGlnAla 379
Db      1173 GTGACTCGGCGCCCGACAGCTGCTGCTACTG-----CACAGCCCATACACACTCT 1226
Qy      380 LeuSer 381
Db      1227 CCCAGC 1232

```

RESULT 9

US-09-614-124B-803

Sequence 803, Application US/09614124B

Patent No. 6630574

GENERAL INFORMATION:

APPLICANT: Wang, Tonglong

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary

APPLICANT: Vedelick, Tom

APPLICANT: Carter, Darick

APPLICANT: Retter, Marc

APPLICANT: Mennion, Jane

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C9

CURRENT APPLICATION NUMBER: US/09/614,124B

CURRENT FILING DATE: 2001-07-11

NUMBER OF SEQ. ID NOS: 1668

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 803

LENGTH: 1238

TYPE: DNA

ORGANISM: Homo sapien

US-09-614-124B-803

```

Best Local Similarity: 31.41% Mismatches: 143
Query Match: 13.25% Indels: 46
DB: 4 Gaps: 13
US-10-029-345A-109 (1-665) x US-09-614-124B-803 (1-1238)
Qy      19 LeuGluSerGlyThrGluValLeuLeuIleAspSerArgProPheValGluThrAsn 38
Db      168 CTCGCCAGCGGC---GGCAAGTCCCTGCTGCTGATGATGACAGACCTTCTCCGGCGCACAGC 224
Qy      39 ThrSerHisIleLeuGluAlaIleAsnIleAsnCysSerIleuLeuValArgIleu 58
Db      225 GCGGGCTACATCTTACAGTTGCTGCAAGCGGTGTAACCACTGTCGCGCGCGG---281
Qy      59 GlnGlnAspIleValLeuIleThrGluLeuIle-----GlnHisSerAlaValHisIle 76
Db      282 GCTAAGAGGCTCCGTCGAGCTGAGAGATCTGCGCCGCGAGAGAGAGATACCGCGCCGC 341
Qy      77 ValAspIleAspCysSerGluValValValIleValIleValIleValIleValIle 96
Db      342 TTCGCTCCGCGCTTACTCGCGGTGATCATGCTTACAGAGAGCGAGCCCGCGCGTGA 401
Qy      97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyValLeuGluValSerPhe--- 115
Db      402 AGCTCCGCGAGAGACAGACACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 461
Qy      116 ---AsnSerValHisLeuLeuAlaGlyIlePheAlaGluPheSerArgCysPheProGly 134
Db      462 CGCACCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 521
Qy      135 LeuCysGluGlyValSerThrIle-----ValProThrCysIleSer 148
Db      522 TTCTGTTCTTAAACCAAGGCTGCGAGCATCCACCCCGGTTCCCGACGCGCAC 561
Qy      149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db      582 GAGCCCTGGAGCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 641
Qy      160 ThrArgIleLeuProAsnLeuIleValGlyCysGlnArgAspValLeuAsnIleValLeu 179
Db      642 GTGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 701
Qy      180 IleGlnAlaAsnGlyIleGlyIleValIleAsnAlaSerIleThrCysProIleProAsp 199
Db      702 CTGAGACCGCTGGGACATCAAGGCTGTGATGTCCTCCGAGACTGCCAAAC---CAC 758
Qy      200 PheIIeProgluSerHisPheLeuArgValProValAsnAspSerPheCysGluValIle 219
Db      759 TTGGAAGACACTATCACTGACAGTGCATCCAGTGAAGATTAACCAAGCGCCGACATC 818
Qy      220 LeuProTfPheAspIleAspSerValAspPheIIeGluValAlaValAsnAsnGlyCys 239
Db      819 AGCTCTGTTGATGAGAGCCATAGAGTACATGATCCCTGAAGAGACTGCCGCGGCGC 878
Qy      240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIle 259
Db      879 GTGTGTGTGACACTCCAGGCGGCGCATCTCGCGGTCCGACCAACATCTGCTCCGCTACCTG 938
Qy      260 MetLeuArgMetAspMetSerLeuAspGluAlaIleArgPheValIleGluValArgPro 279
Db      939 ATGATGAAGAACCGGAGCTGAGGCTGAGAGAGGCTTCGAGTTCGTTAAGCAGCGCCGAGC 998
Qy      280 ThrIleSerProAsnPheAsnPheLeuGlyGluLeuAspIleArgIleValIleVal 299
Db      999 ATCATCTCGCCCAACTTCAAGCTTCATGGGAGCTGCTGAGTTCGAGTCCCAAGTGTG 1058
Qy      300 AsnGlnThrGlyAlaSerGlyProIleSerIleValLeuValIleGluValPro 319
Db      1059 GCCACGCTCTGTGCTGCGGAGGCTGCTAGC-----CCC 1091
Qy      320 AsnGluProValProAlaValSerGluGlyGluIleValSerGluThrProLeuSerPro 339
Db      1092 TCGGAGACCCCTG-----GGGAGAGCGGGCAAGACCCCGCGCACCC 1133

```

```
OY 340 ProCysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSer 359
Db 1134 -----ACCTGCGAGTTCCTTCAAGCTTTTCGGATC---TCCGTGGGC 1172
OY 360 ValProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAla 379
Db 1173 GTGCACTCGGCCCCCAGACAGCTGCTTACTCTG-----CACAGCCCCATCACCACTCT 1226
OY 380 LeuSer 381
Db 1227 CCCAGC 1232

RESULT 10
US-09-671-325-803
; Sequence 803, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongrong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: DIAGNOSIS OF LUNG CANCER
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 803
; LENGTH: 1238
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-803

Alignment Scores:
Pred. No.: 2,14e-37 Length: 1238
Score: 453.00 Matches: 120
Percent Similarity: 50.52% Conservative: 73
Best Local Similarity: 31.41% Mismatches: 143
Query Match: 13.25% Indels: 46
DB: 4 Gaps: 13

US-10-029-345A-109 (1-665) x US-09-671-325-803 (1-1238)
OY 19 LeuGluSerGlyThrGluValLeuLeuLeuAspSerArgProPheValGluTyrAsn 38
Db 168 CTGCCGACGGGC---GGCAAGTGCCTGCTGCTGACAGCTGACAGCCGTTCTCGGCGCACAGC 224
OY 39 ThrSerHisLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 58
Db 225 GCGGAGCTACATCTGAGTTCGATCAAGTGCCTGACACCAATCGGCGGCGGCGG--- 281
OY 59 GlnGlnAspLysValLeuIleThrGluLeuIle-----GlnHisSerAlaLysHisLys 76
Db 282 GCTAAGGAGCTCGGTGAGCTGAGACAGATCTGCCCCCGGAGAGAGAGTACGGGCCCCG 341
OY 77 ValAspIleAspCysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAla 96
Db 342 TTGGGCTCGGCTCTACTCGGCGGTCTATCGTCTACGACGACGCGCGCGCGCGAG 401
OY 97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlnLysLeuGlnLysSerPhe--- 115
Db 402 AGCTTCGCGGAGAGACAGACCGGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 461
OY 116 ---AsnSerValHisLeuLeuValGlyGlyPheAlaGluPheSerArgCysPheProGly 134
Db 462 CGCAGCAGATCTGCTGCTCAAGGCGGCTATGAGAGGTTTCTCCGAGTACCCAGAA 521
```

```
OY 135 LeuCysGluGlyLysSerThrLeu-----ValProThrCysHisSer 148
Db 522 TTCTGTTCTAAACCAAGAGCCCTGAGACCATCCACCCCGGTTCCCGGAGGCGACA 581
OY 149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db 582 GAGCCCTTGAAGCTGAGCTGAGCTGCTGAGGAGCCCACTACACAGACGAGAGGAGCTT 641
OY 160 ThrArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 179
Db 642 GTGAGATCTCTTCCCTCTCTACTGCGGAGAGGCTTACCATGCTGCGGAGAGACATG 701
OY 180 IleGlnHisAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAsp 199
Db 702 CTGAGCGCCCTGGGACATCACGGCTGTGTGAATGTCTCTCGGACGTCGCCAAC---CAC 758
OY 200 PheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle 219
Db 759 TTTCAGAGACATATCATGATCAAGTGCATCCCATGGAAGATTAACCAAGGCCGACATC 818
OY 220 LeuProTrioLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCys 239
Db 819 AGCTCCTGGTTTCATGGAAGCCATAGATCATGATGCGGTGAAGGAGCTGCGGGGCC 878
OY 240 ValLeuValHisCysLeuAlaGlyIleSerArgSerHisThrIleAlaIleAlaTyrIle 259
Db 879 GTGCTGTGACATGCGAGGCGGATCTCGGCTCGGCGCACCATCTGCTGCTTACTCTG 938
OY 260 MetLysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGluLysArgPro 279
Db 939 ATGATGAAGAAACGGGTGAGGCTGAGAGAGGCTTGAATTTGTTAAGCAGCGCCGAC 998
OY 280 ThrIleSerProAsnPheAsnPheLeuGlnLeuLeuAspTyrGluLysIleLys 299
Db 999 ATCATCTCGCCCAACTTCAGCTTCATGAGGCGGACGCTGCTCAGTTCGAGTCCGACG 1058
OY 300 AsnGlnThrGlyValAspSerGlyProLysSerLysLeuLeuHisLeuGluLysPro 319
Db 1059 GCCACGCTCTGTGCTCGGAGGCTGCTAC-----CCC 1091
OY 320 AsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerPro 339
Db 1092 TCGGACCCCTG-----GGGAGAGCGGAGGAGACCCCGCCACCCCC 1133
OY 340 ProCysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSer 359
Db 1134 -----ACCTGCGAGTTCCTTCAAGCTTTCCGATC---TCCGTGGGC 1172
OY 360 ValProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAla 379
Db 1173 GTGCACTCGGCCCCCAGACAGCTGCTTACTCTG-----CACAGCCCCATCACCACTCT 1226
OY 380 LeuSer 381
Db 1227 CCCAGC 1232

RESULT 11
US-09-589-184-803
; Sequence 803, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongrong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C8
```



```

: CURRENT APPLICATION NUMBER: US/09/589,184
: CURRENT FILING DATE: 2000-06-05
: NUMBER OF SEQ ID NOS: 827
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 803
: LENGTH: 1238
: TYPE: DNA
: ORGANISM: Homo sapien
US-09-589-184-803

Alignment Scores:
Pred. No.:      2,14e-37      Length:      1238
Score:          453.00       Matches:     120
Percent Similarity: 50.52%    Conservative: 73
Best Local Similarity: 31.41% Mismatches:   143
Query Match:     13.25%      Indels:      46
DB:              Gaps:       13

US-10-029-345A-109 (1-665) x US-09-589-184-803 (1-1238)

QY      19 LeuGlusertGlyThrGluValLeuLeuLleApsSerArgProPheValGluTyAsn 38
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dd      168 CTGCCAGCGGGC---GCAGAATGCTCTGTGCTGACAGCAACCGTTCCTTGCCACAGC 224
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      39 ThSerHietIleuLnuAlalIleAnlIleAnCySserLybleumetyAtgAdglou 58
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      225 GCGGGCTACACTCTAGATTCTGGTCAACGCGCTGTAAACAACATCTGCGGGCGG-- 281
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      59 GlnglnAapryValIleuIleThrgluLeulle-----GlnHlseryAlalyHlelys 76
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      282 GCtAAAGGCTCCTGTAGAGCTGTGAAGATCTCCGCCGCGAGAGAGAGTACGGCCGC 341
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      77 ValAapLIleApCySserGlnyValValValTYAapGInSerSerGIaapvalAla 96
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      342 TTCCGCTCCGCGCTTACTCGCGGCTCATCGTTCAAGACAGCGACGCCGCCGCCGAG 401
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      97 SetIeusertSerApCyPheleuthrValIleuLeuglylybleuGluysertPhe--- 115
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dd      402 AGCTCTCGCGAGGACAGCACCGTGTGCGTGGTGGAGCGCGCTCGCGCGAACGCCGAG 461
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      116 ---AmnseryAlalIspeuLeuaIaglyIlePhelaglupheserArgCyPheProgly 134
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      462 CGCAACGACATCTGCTGTCTCAAAGCGGCTATGAGAGTTTCTCCGAATACCAGA 521
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      135 leuCysgluGlylyaserThrieu-----ValProthrCysAliser 148
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dd      522 TTCGTGTTTAAMCAAGGCCCTGGAGCAATCCACCCCGGTTCCCCCAGCGCAC 581
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      149 GlnPro-----CysleuProvalAla-----AmnlleglyPro 155
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      582 GAAGCCTTGGACCTGTGACCTGACGCTCTGTGGAGACCCCAACAGACAGAGAGGTCCT 641
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      160 ThrArgIleuProAnleuTyTrileuGlyCywglnatgarpyValIleuanlyGlu 179
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dd      642 GTGGAGATCTTCTCTCTCTCTACCTCGGCAAGTGCCTACATGCTGCCGAGAGACATG 701
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      180 IleglnglnAnanglylleGlyTyValIleuAnaIasertyThrCywProlysProAps 199
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      702 CTGAGCGCCCTGGGACATCAACGGCTGTTGAATGTCTCTGGAAGTCCCAAC---CAC 758
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      200 PheIleProGluSerHiePheleuarGyalProvalAamApsSerPheCysgluyle 219
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      759 TTTGAAGACACATATCAAGTACCAAGTCCCAAGTAGAAGTAAACACAGAGCGCAATC 818
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      220 leuProthrIleuAaprylservAlaapPheIIegulyvalAlayAlaserAnglyCys 239
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dd      819 AGCTCTTGTTCAATGGAAGCCATAGATCATCTGATCCGTGAAGAACTGCCCTGGGGCC 878
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      240 ValIeuValHiasCyblenuIaglyIleseratgserrathrlleaIIeaIatyIle 255
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dd      879 GTGTGTGTGATGTCACAGGCGGACATCTCGCGGTGGCCACATCTGCTGGCTACCTG 938
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      260 MetIyArGmetAapMetserIleuAapGluAlatyArgPheVallyGluIyArGPro 279

```

```

Db      939 ATGATGAAGAAACGGGTGAGCTGAGAGAGCGCTTCGAGTTCCGTAAAGCAGCCGCCAGC   998
           |||          ::::          |||          |||          |||          |||
Qy      280 ThrIleSerProAsnPheAnPheLeuGlyGlnLeuLeuAspYrGluValysIleLys   299
           |||||         |::::|        |::::|       |::::|      |::::|
Db      999 ATCATCTCGGCCCAATTCACTTCACTTGAGGCAGCTGCTGAGTTCAGTCCAGAGTCTG   1058
           |||||         |::::|        |::::|       |::::|      |::::|
Qy      300 AsnGlnThrGlyAlaSerGlyProLysSerTyrLeuLysLeuHisLeuGlnLysPro   319
           |||          ::::          |||          |||          |||          |||
Db      1059 GCCACGTCCTGCTGCTGCGGAGGCTGCTAGC-----CCC   1091
           |||          ::::          |||          |||          |||          |||
Qy      320 AsnGlnProValProAlaAlaValSerGlnGlyGlnLysSerGlnThrProLeuSerPro   339
           ::|          |||          |||          |||          |||          |||
Db      1092 TCGGGAACCCCTG-----GGGGAGCGGGGCAAGACCCTCCGACACCCCC   1133
           |||          |||          |||          |||          |||          |||
Qy      340 ProCysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSer   359
           |||||         |::::|        |::::|       |::::|      |::::|
Db      1134 -----ACCTCGCAGCTTCGCTTCACGCTTCCGATC---TCGGTGGGC   1172
           |||||         |::::|        |::::|       |::::|      |::::|
Qy      360 ValProSerValProSerValGlnProSerLeuGlnAspProLeuValGlnAla   379
           |||||         |::::|        |::::|       |::::|      |::::|
Db      1173 GTGACTCGGCCCCCGACGACCTGCTTACTTG----CACAGCCCATCACCACTCT   1226
           |||||         |::::|        |::::|       |::::|      |::::|
Qy      380 LeuSer 381
           |||
Db      1227 CCCAGC 1232

RESULT 12
US-09-016-434-1100
; Sequence 1100, Application US/09016434
; Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Sellhammer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1100:
SEQUENCE CHARACTERISTICS:
LENGTH: 2240 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1255784

```



```

Db      745 AGCTTCGGAGAGACGACCGGCTGCTGTGTGTGAGGCGGTGGCGGCAACGGCAG 804
Qy      116 ---AanSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGly 134
Db      805 CGCACCAGACATCTGCTGCTCAAAAGCGGCTAAGAGGTTTCTCCGAGTACCCAGAA 864
Qy      135 LeuCyGluGluGlySerThrLeu-----ValProThrCysIleSer 148
Db      865 TTCTGTTCTTAAACCAAGGCCCTGGAGCCATCCACCCCGGTCTCCCGCACTGCCACA 924
Qy      149 GlnPro-----Cys---LeuProValAlaAsnIle---GlyPro 159
Db      925 GAGCCCTTGGACCTGGGCTGCACTCTCTGTGGAGCCCACTACACACAGGGGGGTCT 984
Qy      160 ThrArgIleLeuProAsnLeuTyLeuGlyCysGlnArgAspValLeuAsnIleGluLeu 179
Db      985 GTGGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1044
Qy      180 ILeGlnAlaAsnGlyIleGlyTyValLeuAsnAlaSerTyThrCysProIleProAsp 199
Db      1045 CTGGAGCCCTGGGCTGAGCATCGCTCTGTGAATGCTCTCTGGAGCTGCCAACC---CAC 1101
Qy      200 PheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIle 219
Db      1102 TTGGAAGACACTATCATGATCAAGTCATCCAGTGGAGATTAACCAAGCCGACATC 1161
Qy      220 LeuProTyrLeuAspIleSerValAspPheIleGluValAlaIleValAsnAsnGlyCys 239
Db      1162 AGCTCTGCTTCAATGAGACCATAGATCATGATCATGATCATGATCATGATCATGATCATG 1221
Qy      240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyIle 259
Db      1222 GTGCTGTGACTGACCGACGGGCGCATCTCGCGGTCCGACCAATCTGCTGGCTTACCTG 1281
Qy      260 MetLeuArgMetAspMetSerLeuAspGluAlaTyTrpAspValIleGlyIleGlyIleGly 279
Db      1282 ATGATGTAAGAAAGGGGTGAGAGCTGGAGAGCCCTTGCAGTCTTGAAGACGCGCCGAC 1341
Qy      280 ThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyGluIleIleIle 299
Db      1342 ATCATCTCGCCCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1401
Qy      300 AsnGlnThrGlyIleSerGlyProIleSerIleLeuIleLeuIleIleGluIleIle 319
Db      1402 GCCACGTCCTGCTGCTGAGGCTGTAGC-----CCC 1434
Qy      320 AsnGluProValProAlaValSerGluGlyGlnIleSerGluThrProLeuSer---Pr 339
Db      1435 TCGGAGACCTCTG-----CGGAGCGGGGCAAGACCCCGCCACCC 1476
Qy      339 oProCysAlaAspSerAlaThr-----SerGluAlaAlaGlyGlnArgProValHisPr 357
Db      1477 ACCTCCAGATGTCCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1530
Qy      357 oAlaSerValProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuVa 377
Db      1531 AGCAGCTGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1578
Qy      377 ILeGlnAlaLeuSerGly 382
Db      1579 TAGAGCCGCTGGGG 1594

```

```

; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Ligu
; APPLICANT: Wang, Aljun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 825
; LENGTH: 2064
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-825

Alignment Scores:
Pred. No.: 1 02e-36 Length: 2064
Score: 450.50 Matches: 122
Percent Similarity: 51.81% Conservative: 78
Best Local Similarity: 31.61% Mismatches: 140
Query Match: 13.18% Indels: 46
DB: 4 Gaps: 14

US-10-029-345A-109 (1-665) x US-09-736-457-825 (1-2064)
Qy      19 LeuGluSerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyAsn 38
Db      511 CTGCCAGCGGC---GGCAAGTGCCTGCTGTGATGTCGATGCAAGACCTTCTCGGCGCAGC 567
Qy      39 ThrSerHisIleLeuGluAlaIleAsnIleAsnCysSerIleLeuMetIleAsnArgLeu 58
Db      568 GCGGGCTATCTCTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 624
Qy      59 GlnGlnAspIleValLeuIleThrGluLeuIle-----GlnHisSerAlaIleValHisIle 76
Db      625 GCTAAGGCTCTGCTGAGCTGAGCATCTTCCCGCGAGAGAGAGAGATCGCCCGC 664
Qy      77 ValAspIleAspCysSerGlnIleValValValTyAspGlnSerGlnIleValAla 96
Db      685 TTGCGCTCGGCTTACTCGGCGGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 744
Qy      97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyIleLeuGluIleSerPhe--- 115
Db      745 AGCTTCGGAGAGACGACCGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 804
Qy      116 ---AanSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGly 134
Db      805 CGCACCAGACATCTGCTGCTCAAAAGCGGCTAAGAGGTTTCTCCGAGTACCCAGAA 864
Qy      135 LeuCyGluGluGlySerThrLeu-----ValProThrCysIleSer 148
Db      865 TTCTGTTCTTAAACCAAGGCCCTGGAGCCATCCACCCCGGTCTCCCGCACTGCCACA 924
Qy      149 GlnPro-----Cys---LeuProValAlaAsnIle---GlyPro 159
Db      925 GAGCCCTTGGACCTGGGCTGCACTCTCTGTGGAGCCCACTACACACAGGGGGGTCT 984
Qy      160 ThrArgIleLeuProAsnLeuTyLeuGlyCysGlnArgAspValLeuAsnIleGluLeu 179
Db      985 GTGGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1044
Qy      180 ILeGlnAlaAsnGlyIleGlyTyValLeuAsnAlaSerTyThrCysProIleProAsp 199
Db      1045 CTGGAGCCCTGGGCTGAGCATCGCTCTGTGAATGCTCTCTGGAGCTGCCAACC---CAC 1101
Qy      200 PheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIle 219
Db      1102 TTGGAAGACACTATCATGATCAAGTCATCCAGTGGAGATTAACCAAGCCGACATC 1161
Qy      220 LeuProTyrLeuAspIleSerValAspPheIleGluValAlaIleValAsnAsnGlyCys 239

```


Qy 377 |G|A|A|LeuserG|Y 382
|::|||
Db 1579 TAGAGCCGCCCTGGGG 1594

Search completed: June 22, 2004, 02:46:14
Job time : 150.037 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:13:37 ; Search time 14.4416 Seconds
(without alignments)
2397.707 Million cell updates/sec

Title: US-10-029-345A-109
Sequence: 1 MAHEMIGTQIVTERIVALLE.....LGKVGSSSPSSGMEIIEVVS 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3406	99.6	665	1 DUSG_HUMAN	Q9BY84 homo sapien
2	1326	38.8	625	1 DUS8_HUMAN	Q13202 homo sapien
3	1302	38.1	663	1 DUS8_MOUSE	009112 mus musculu
4	497.5	14.6	619	1 VHP1_CAEEL	010038 caenorhabdi
5	471.5	13.8	381	1 DUS6_HUMAN	Q16828 homo sapien
6	469	13.7	482	1 DUS6_MOUSE	Q9566 mus musculu
7	468	13.7	483	1 DUS6_MOUSE	Q9566 mus musculu
8	466.5	13.6	381	1 DUS6_MOUSE	Q9566 mus musculu
9	465.5	13.6	381	1 DUS6_MOUSE	Q9566 mus musculu
10	445	13.0	394	1 DUS4_HUMAN	Q13115 homo sapien
11	444	13.0	375	1 DUS4_HUMAN	Q9PW71 gallus gall
12	439.5	12.9	395	1 DUS4_MOUSE	Q62767 rattus norv
13	433	12.7	367	1 DUS1_MOUSE	P28563 mus musculu
14	423	12.4	367	1 DUS1_MOUSE	P28563 mus musculu
15	420	12.3	367	1 DUS1_MOUSE	P28563 mus musculu
16	418	12.2	367	1 DUS1_MOUSE	P28563 mus musculu
17	417	12.2	367	1 DUS1_MOUSE	P28563 mus musculu
18	409	12.0	318	1 DUS2_HUMAN	Q05923 mus musculu
19	390	11.4	320	1 DUS7_HUMAN	Q16829 mus sapien
20	386.5	11.3	320	1 DUS7_MOUSE	Q1426 mus musculu
21	367	10.7	384	1 DUS7_MOUSE	Q54838 rattus norv
22	361.5	10.6	280	1 DUS7_MOUSE	Q54838 rattus norv
23	360.5	10.5	280	1 DUS7_MOUSE	Q54838 rattus norv
24	269	7.9	198	1 DUS5_HUMAN	Q16630 homo sapien
25	268	7.8	198	1 DUS5_HUMAN	Q16630 homo sapien
26	205.5	6.0	188	1 DUS1_MOUSE	Q95147 mus musculu
27	205.5	6.0	217	1 DUS1_MOUSE	Q95147 mus musculu
28	205.5	6.0	223	1 STYX_MOUSE	Q8WU12 homo sapien
29	205	6.0	223	1 STYX_MOUSE	Q8WU12 homo sapien
30	204	6.0	278	1 FMPI_YEAST	Q09099 mus musculu
31	202	5.9	364	1 FMPI_YEAST	Q09099 mus musculu
32	201.5	5.9	313	1 MSTY_HUMAN	Q12453 saccharomyc
33	196	5.7	185	1 DUS3_HUMAN	Q9Y618 homo sapien
					P51452 mus musculu

ALIGNMENTS

RESULT 1	DUSG_HUMAN	STANDARD;	PRT;	665 AA.
AC	Q9BY84; Q9COG3;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Dual specificity protein phosphatase 16 (EC 3.1.3.48) (EC 3.1.3.16)			
DE	(Mitogen-activated protein kinase phosphatase 7) (MAP kinase phosphatase 7) (MKP-7).			
GN	DUSP16 OR MKP7 OR KIAA1700.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21486429; PubMed=11489891;			
RA	Mapada K., Shima H., Watanabe M., Kikuchi K.;			
RT	"MKP-7, a novel mitogen-activated protein kinase phosphatase, functions as a shuttle protein."			
RT	J. Biol. Chem. 276:39002-39011(2001).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Brain;			
RC	MEDLINE=21082932; PubMed=11214970;			
RX	Ngase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;			
RA	"Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."			
RT	DNA Res. 7:347-355(2000).			
RL	DNA Res. 7:347-355(2000).			
CC	- FUNCTION: Involved in the inactivation of MAP kinases.			
CC	- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein tyrosine + phosphate.			
CC	- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein + phosphate.			
CC	- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.			
CC	- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.			
CC	- SIMILARITY: Contains 1 rhodanese domain.			
CC	- SIMILARITY: Contains 1 rhodanese domain.			
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL: AB052156; BAB40814.1; -			
CC	EMBL: AB051487; BAB21791.1; ALT_INIT.			
CC	HSSP: Q16828; IMKP.			
CC	Genew; HGNC:17909; DUSP16.			
CC	MIM; 607175; -			
CC	GO; GO:0005737; C:cytoplasm; TAS.			
CC	GO; GO:0005634; C:nucleus; TAS.			

34	194.5	5.7	185	1 DUS3_MOUSE	Q9d7x3 mus musculu
35	189	5.5	489	1 MS65_YEAST	P38590 saccharomyc
36	185	5.4	807	1 YBPT_YEAST	P38148 saccharomyc
37	183.5	5.4	276	1 PTP3_CHIEU	Q39491 chlamydomon
38	181.5	5.3	295	1 DUSF_HUMAN	Q9H122 mus musculu
39	177	5.2	125	1 DUSF_MOUSE	Q81472 mus musculu
40	171.5	5.0	339	1 DUSC_MOUSE	Q9d0c2 mus musculu
41	170.5	5.0	292	1 Y042_CAEEL	P34680 caenorhabdi
42	161	4.7	198	1 DUSC_HUMAN	Q9u116 mus musculu
43	158.5	4.6	340	1 DUSC_HUMAN	Q9u116 mus musculu
44	157	4.6	209	1 Y112_YEAST	P40479 saccharomyc
45	156	4.6	198	1 DUSD_MOUSE	Q9y17 mus musculu

DR GO; GO:0004721; F:protein phosphatase activity; TAS.
 DR GO; GO:0016311; P:dephosphorylation; TAS.
 DR GO; GO:0000188; P:inactivation of MAPK; TAS.
 DR GO; GO:0045209; P:leptocystin B-sensitive MAPK phosphatase nuc. . .; TAS.
 DR GO; GO:0045204; F:MAPK nucleus export; TAS.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANESE_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolase; Nuclear protein.
 FT DOMAIN 22 137 RHODANESE.
 FT DOMAIN 227 289 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 244 244 PHOSPHOCYSTEINE INTERMEDIATE (BY SIMILARITY)
 SQ SEQUENCE 665 AA; 73101 MW; 1BD853FF08460DF CRC64;

Query Match 99.6%; Score 3406; DB 1; Length 665;
 Best Local Similarity 99.7%; Pred. No. 7.7e-200;
 Matches 665; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALESGTEKYLIDSRPEVYNTSHLEAININCSGLMRRIQQ 60
 DB 1 MAHEMIGTQVTERLVALESGTEKYLIDSRPEVYNTSHLEAININCSGLMRRIQQ 60
 QY 61 DKYLITELLQHSAKHKVDICQKVYVYDSSQVASSDCEFLVLLGKLEKSNVHL 120
 DB 61 DKYLITELLQHSAKHKVDICQKVYVYDSSQVASSDCEFLVLLGKLEKSNVHL 120
 QY 121 LAGGAFSFRCPGCEGKSTVPTTCISQPCPVANIGPTILLPNLYGCGRDVANKELI 180
 DB 121 LAGGAFSFRCPGCEGKSTVPTTCISQPCPVANIGPTILLPNLYGCGRDVANKELI 180
 QY 121 LAGGAFSFRCPGCEGKSTVPTTCISQPCPVANIGPTILLPNLYGCGRDVANKELI 180
 DB 121 LAGGAFSFRCPGCEGKSTVPTTCISQPCPVANIGPTILLPNLYGCGRDVANKELI 180
 QY 181 QONGIGYLVNASTCPKDPFIPESHFLRPVNDSCFKILPMWKSVDPIEKARNGCV 240
 DB 181 QONGIGYLVNASTCPKDPFIPESHFLRPVNDSCFKILPMWKSVDPIEKARNGCV 240
 QY 181 QONGIGYLVNASTCPKDPFIPESHFLRPVNDSCFKILPMWKSVDPIEKARNGCV 240
 DB 181 QONGIGYLVNASTCPKDPFIPESHFLRPVNDSCFKILPMWKSVDPIEKARNGCV 240
 QY 241 LVHCLAGTSRSATTAIATYMKRMDSLDEAYRFVYKRPPTISPNFNLGQLLDYEKKIK 300
 DB 241 LVHCLAGTSRSATTAIATYMKRMDSLDEAYRFVYKRPPTISPNFNLGQLLDYEKKIK 300
 QY 301 QTGAAGPKSKLKLHLLEKNEPVPVAVSGGQKSTPLSPCADSATSEAGQRPVHPASV 360
 DB 301 QTGAAGPKSKLKLHLLEKNEPVPVAVSGGQKSTPLSPCADSATSEAGQRPVHPASV 360
 QY 361 PSVSVVQSLLEDSPVVALSGLHLISADRLSDSKLKRSPSLDIKSYVSASMAASLHGF 420
 DB 361 PSVSVVQSLLEDSPVVALSGLHLISADRLSDSKLKRSPSLDIKSYVSASMAASLHGF 420
 QY 421 SSSSDALEYKPSFTLLDGTNKLCOFSPVOELSPOTPTSTDKERASIPKKLQTPARPSDQ 480
 DB 421 SSSSDALEYKPSFTLLDGTNKLCOFSPVOELSPOTPTSTDKERASIPKKLQTPARPSDQ 480
 QY 481 SKRLHVSRTSSSGTAQSRSLSPILRSGSVEDNYHTSFLFGLSTSQOHLTKSAGLGLKGMH 540
 DB 481 SKRLHVSRTSSSGTAQSRSLSPILRSGSVEDNYHTSFLFGLSTSQOHLTKSAGLGLKGMH 540
 QY 541 SDLIAPGSTPSSLSSVYFATSSSHPTSAATYGSASISYSCQLPTCGDQVYSVRRR 600
 DB 541 SDLIAPGSTPSSLSSVYFATSSSHPTSAATYGSASISYSCQLPTCGDQVYSVRRR 600
 QY 601 QKPSDRADSRSSWHEESPFEKQPKRRSCQWFGESIMSENRSEBELGKYGSSSFGSGWE 660
 DB 601 QKPSDRADSRSSWHEESPFEKQPKRRSCQWFGESIMSENRSEBELGKYGSSSFGSGWE 660

QY 661 IIEVS 665
 DB 661 IIEVS 665

RESULT 2

DUS8_HUMAN
 ID DUS8 HUMAN STANDARD; PRT; 625 AA.
 AC Q13202;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Dual specificity protein phosphatase hvh-5).
 GN DUS8 OR VH5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=96009533; PubMed=7561881;
 RA Martell K.J., Seasholtz A.F., Kaak S.P., Clemens K.K., Dixon J.E.;
 RT "hvh-5: a protein tyrosine phosphatase abundant in brain that
 RL inactivates mitogen-activated protein kinase.";
 J. Neurochem. 65:1823-1833(1995).
 CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
 CC phosphatase as well as with serine/threonine-protein phosphatase (by
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- TISSUE SPECIFICITY: Abundant in brain, heart and skeletal muscle.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; U27193; AA83151.1; -.
 DR HSSP; Q16828; IMPR.
 DR Genew; HGNC:3074; DUSP8.
 DR MIM; 602038; -.
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0004725; F:protein tyrosine phosphatase activity; TAS.
 DR GO; GO:0000188; P:inactivation of MAPK; TAS.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANESE_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolase; Nuclear protein.
 FT DOMAIN 23 138 RHODANESE.
 FT DOMAIN 162 430 PROTEIN-TYROSINE PHOSPHATASE.
 FT DOMAIN 310 550 PRO-RICH.
 FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY SIMILARITY).

SQ SEQUENCE 625 AA; 65840 MW; DOBEAL44487219666 CRC64;
 Query Match 38.8%; Score 1326; DB 1; Length 625;
 Best Local Similarity 45.1%; Pred. No. 2,4e-73;
 Matches 314; Conservative 94; Mismatches 185; Indels 104; Gaps 21;
 QY 1 MAHEMIGTOIV-TERIVALLESGTEKYLIDSRPFVYNTSHILEAININCSKLMRRRLQ 59
 DB 1 MAGDRIPRKVMADKLASLLRGPGPGLVDSRPFVYNSCHVLSVNICSLTVKRRRLQ 60
 QY 60 QDKVLTTELIIQSAKHKVDIDCSQKVVYVYDQSSQDVASLSSDCFLTVLGLKLEKSPSVH 119
 DB 61 QGKVTIAELIQPARSGVEATEBPQDVVYDQSTRDASVLAADSPSLISKLDCGCFDSVA 120
 QY 120 LLAGFAEBSRCRPGCEGK-STLVPTICISQPLPVANITPTRLPLNLYGCGQDVLANKE 178
 DB 121 ILTGFAFATSSCPGCEGKPAALPMSLSQPCLPVPSVGLTRILPLNLYGSGQDVLANKE 180
 QY 179 LIOONGIGVYLANASYCPRKDPFIPESHFLRVPNVDSCEKILPWLKSDVFIEKAKASNG 238
 DB 181 LMTQNGISVYLANNSCPKDPFICESHFRVPIINDYCEKILPWLKSDVFIEKAKASNG 240
 QY 239 CVLVHCLAGISRSATTAIAYIMKMDSLDEAYRVEKRPITSPENFLGQLLDYEKKI 298
 DB 241 QVIVHCLAGISRSATTAIAYIMKMGSSDADRVFVDRRPSISPNENFLGQLLDYEKKI 300
 QY 299 KNOTGASGPKSKLKLHLEKPN--EVPVAVSEGGQKSETPPLPCADSAT--SEAAQR 353
 DB 301 KLALALQGDG-----TPSGTPEPPPPAPAGAPLRPLP--PPTSSEATGNAAAREG 350
 QY 354 PVHAPASVPSVQPSLLEDSPLVQALSGHLSDRLSDNCKLRSPSLIKSVYSASM 413
 DB 351 GLTAGGEPAPPPPPA--TSALQQGRGLHLSDRLODTRNLKRSVSLDKA----- 401
 QY 414 AASLHGSSSEDALEYYKPTTLDTN-----KLQCP--SP--VQLESEOTPEPSPD 461
 DB 402 -----YAPRRRDPGPPPPPPGPAKCKLSDSGALGSSSPD--SPD 444
 QY 462 KEBASIPKQLQTPAPSDSQKRLHVSRTSSGTAQKSLPLHRSQVSDNHTSFLPG- 520
 DB 445 AADEAPRRRRPRP-----PAGSPAR--SPASLIG-----LNFCD 478
 QY 521 --LSTSQHITKAGIGLKG-----WHDILAQTSPTSLTSSWYFATESHRYSA 569
 DB 479 AARQTPRHGSLASAPGLPPGQAPAGAWAPPLDSE--GTPPDGWCSPS----- 529
 QY 570 SAIVGGSASVAYSCQLPTCGDQVYVVRROKPSDRDARSWHEESPPEKQPKRRSCQ 629
 DB 530 GAGAGAGVLFAPGRAGAPRGSGSLRRREAAAEPRDARTGMPPEPAPETQPKRRSCQ 589
 QY 630 MEFGEISMSNSR-BELGKVGQSQSFSGMEIIEVS 665
 DB 590 MEFEBEG-MVEGRARGBELALGKQSPSGSVEVEVS 625
 RESULT 3
 DUS8_MOUSE STANDARD; PRT; 663 AA.
 ID DUS8_MOUSE
 AC 009112;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
 GN (Neuronal tyrosine threonine phosphatase 1).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96311565; PubMed=8733137;

RA Theodosiou A.M., Rodrigues N.R., Nesbit M.A., Ambrose H.J.,
 RA Paterson H., McEldan Arnold E., Boyd Y., Leverhna M.A., Owen N.,
 RA Blake D.J., Ashworth A., Davies K.E.,
 RT "A member of the MAP kinase phosphatase gene family in mouse
 RT containing a complex trinucleotide repeat in the coding region,"
 RL Hum. Mol. Genet. 5:675-684(1996).
 CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
 CC phosphate as well as with serine/threonine-protein phosphate (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
 CC -1- TISSUE SPECIFICITY: Expressed predominantly in brain and lung.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; X95518; CAA64772.1; -;
 DR HSSP; Q16828; IMKP.
 DR MGD; MG1.106626; Dusp8.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; Dspc; 1.
 DR PRINTS; PR01764; MAPKPHPTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANES_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolyase; Nuclear protein.
 FT DOMAIN 23 138 RHODANES.
 FT DOMAIN 162 432 PROTEIN-TYROSINE PHOSPHATASE.
 FT DOMAIN 452 459 POLY-ARG.
 FT DOMAIN 555 558 POLY-SER.
 FT DOMAIN 576 576 POLY-GLY.
 FT DOMAIN 577 600 POLY-SER.
 FT DOMAIN 311 552 PRO-RICH.
 FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
 FT SIMILARITY).
 SQ SEQUENCE 663 AA; 68847 MW; 416F429A12C1FA7C CRC64;
 Query Match 38.1%; Score 1302; DB 1; Length 663;
 Best Local Similarity 43.4%; Pred. No. 7,4e-72;
 Matches 318; Conservative 102; Mismatches 176; Indels 136; Gaps 26;
 QY 1 MAHEMIGTOIV-TERIVALLESGTEKYLIDSRPFVYNTSHILEAININCSKLMRRRLQ 59
 DB 1 MAGDRIPRKVMADKLASLLRGPGPGLVDSRPFVYNSCHVLSVNICSLTVKRRRLQ 60
 QY 60 QDKVLTTELIIQSAKHKVDIDCSQKVVYVYDQSSQDVASLSSDCFLTVLGLKLEKSPSVH 119
 DB 61 QGKVTIAELIQPARSGVEATEBPQDVVYDQSTRDASVLAADSPSLISKLDCGCFDSVA 120
 QY 120 LLAGFAEBSRCRPGCEGKSTLVPT-CTSQPLPVANITPTRLPLNLYGCGQDVLANKE 178
 DB 121 ILTGFAFATSSCPGCEGKPAALPMSLSQPCLPVPSVGLTRILPLNLYGSGQDVLANKE 180
 QY 179 LIOONGIGVYLANASYCPRKDPFIPESHFLRVPNVDSCEKILPWLKSDVFIEKAKASNG 238
 DB 181 LMTQNGISVYLANNSCPKDPFICESHFRVPIINDYCEKILPWLKSDVFIEKAKASNG 240

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Forebrain;
 RX MEDLINE=66312959; PubMed=8670865;
 RA Groom L.A., Sneddon A.A., Alessi D.R., Dowd S., Keyse S.M.;
 RT "Differential regulation of the MAP, Sap and Rk/p38 kinases by Pyst1,
 RL a novel cytosolic dual-specificity phosphatase.";
 EMBO J. 15:3621-3632(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Liver;
 RX MEDLINE=99077745; PubMed=9858808;
 RA Furukawa T., Yatsunaka T., Yousef E.M., Abe T., Yokoyama T.,
 RA Fukushima S., Soeda E., Hoshi M., Hayashi Y., Sunamura M., Kobari M.,
 RA Horii A.;
 RT "Genomic analysis of DUSP6, a dual specificity MAP kinase phosphatase,
 RT in pancreatic cancer.";
 RL Cytogenet. Cell Genet. 82:156-159(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT VAL-114.
 RC TISSUE=Colon, Kidney, Skin, and Stomach;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Mantzaris K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
 RA Pahey J., Hailton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Stevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 204-347.
 RX MEDLINE=99140299; PubMed=10048930; McDonald N.Q.;
 RA Stewart A.E., Dowd S., Keyse S.M., McDonald N.Q.;
 RT "Crystal structure of the MAPK phosphatase Pyst1 catalytic domain and
 RT implications for regulated activation.";
 RL Nat. Struct. Biol. 6:174-181(1999).
 CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
 CC family.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q16828-1; Sequence=Displayed;
 CC Name=2; Synonyms=DUSP6-ALT;
 CC IsoId=Q16828-2; Sequence=VSP_005137;
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; X3920; CAA63813.1; -;
 CC DR EMBL; AB013601; BAA31968.1; -;
 CC DR EMBL; AB013383; BAA31968.1; JOINED.
 CC DR EMBL; AB013600; BAA31968.1; JOINED.
 CC DR EMBL; AB013382; BAA31969.1; -;
 CC DR EMBL; AB013602; BAA31969.1; -;
 CC DR EMBL; BC003143; AAH03143.1; -;
 CC DR EMBL; BC003562; AAH03562.1; -;
 CC DR EMBL; BC005047; AAH05047.1; -;
 CC DR EMBL; BC037236; AAH37236.1; -;
 CC DR EMBL; IMKP; 22-JUL-99.
 CC DR PDB; 1H2M; 25-JAN-02.
 CC DR Genew; HGNC:3072; DUSP6.
 CC DR MIM; 602748; -;
 CC DR GO; GO:0005737; Cytoplasm; IDA.
 CC DR GO; GO:0005625; Cytosol; TAS.
 CC DR GO; GO:0004722; F-protein serine/threonine phosphatase activity; IDA.
 CC DR GO; GO:0004725; F-protein tyrosine phosphatase activity; IDA.
 CC DR GO; GO:0000188; P-protein tyrosine phosphatase activity; IDA.
 CC DR GO; GO:0006470; P-protein amino acid dephosphorylation; TAS.
 CC DR InterPro; IPR000340; DS phosphatase.
 CC DR InterPro; IPR008343; MAPK_phosph.
 CC DR InterPro; IPR001763; Rhodanese-like.
 CC DR InterPro; IPR00387; TYR_phosphatase.
 CC DR Pfam; PF00782; DSPC; 1.
 CC DR Pfam; PF00581; Rhodanese; 1.
 CC DR PRINTS; PRO1764; MAPKPHRTASE.
 CC DR SMART; SM00195; DSPC; 1.
 CC DR SMART; SM00450; RHOD; 1.
 CC DR PROSITE; PS50206; RHODANSE_3; 1.
 CC DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
 CC DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 CC DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 CC DR Hydrolase; Polymorphism; Alternative splicing; 3D-structure.
 CC DR DOMAIN; 30 148
 CC DR PROSITE; PS00383; RHODANSE.
 CC DR ACT SITE; 206 381
 CC DR ACT SITE; 293 293
 CC DR VARSPPLIC; 134 279
 CC FT VARIANT; 114 114
 CC FT STRAND; 208 211
 CC FT TURN; 212 213
 CC FT STRAND; 214 217
 CC FT TURN; 219 220
 CC FT HELIX; 225 230
 CC FT TURN; 231 232
 CC FT STRAND; 233 238
 CC FT STRAND; 246 250
 CC FT TURN; 251 252
 CC FT STRAND; 253 257
 CC FT TURN; 254 265
 CC FT HELIX; 269 271
 CC FT HELIX; 272 284
 CC FT TURN; 285 286
 CC FT STRAND; 288 292
 CC FT HELIX; 298 312
 CC FT TURN; 313 315
 CC FT HELIX; 316 326
 CC FT TURN; 328 329
 CC FT TURN; 335 336
 CC FT HELIX; 337 345
 CC FT TURN; 346 346
 CC SQ SEQUENCE; 381 AA; 42333 MW; 03BC12252CE73B26 CRC64;
 CC Query Match 13.8%; Score 471.5; DB 1; Length 381;
 CC Best Local Similarity 34.0%; Pred. No. 1.1e-21;
 CC Matches 113; Conservative 63; Mismatches 107; Gaps 8;

```

QY 19 LESGTEKVLIDSPFVETNTSHILEATININCSKLMKRILOQDKVITTELIOHSA-KHKV 77
DB 27 LEIGMERILLMDCRQGLYESHSHIESAIVNAIPGIMTRRLOKGNLPRALFTGREDRDF 86
QY 78 DIDC-SQKVVYVDSQD-VASLSSDCEFLVTLGLKLEKSPNSVHLIAGGFAEFSRCFPL 135
DB 87 TRRCCTIVVLYDESSDMNNTGEGESILGLIKLKDECRAPVYLEGFSKFOAEFSLH 146
QY 136 CEKSTLVPYC-1SOPCLPVANIG----- 158
DB 147 CE--TNLDGSCSSSPPLVGLGGLRISDSSSDIESLIDRDPNSATDSDSPLSNQ 204
QY 159 --PTRIPLNYLIGQGVVANKELIQONGIGYVINAAYTPKRP-DFIPESHFLRPVND 215
DB 205 SFPYELLIPYLGCAKSTNLDVLEERGIKYLIVNPNLNFENAGFRKQKQIFISHW 264
QY 216 CEKLPWLKSDVDFIEKAKASNGCVLHCLAGSSATAIAYIMKRDMSLDEAYR 275
DB 265 SGNLSQFPFPAISFIDBAGKNGCVLVHCLAGISSVTVVAYVMQKILMSNDAYDI 324
QY 276 EKPTTSPNENFGQLDVEKIKRGTGASGP 307
DB 325 MKKSNISPNENFGQLDERTL-----GLSSP 352

```

RESULT 6

```

DUSA_HUMAN STANDARD; PRT; 482 AA.
AC Q9Y6W6;
DT 28-FEB-2003 (rel. 41, Created)
DT 28-FEB-2003 (rel. 41, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 5) (MAP kinase
DE phosphatase 5) (MKP-5).
GN DUSP10 OR MKP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9932129; PubMed=10391943;
RA Tanoue T., Moriyuchi T., Nishida E.;
RT "Molecular cloning and characterization of a novel dual specificity
RT phosphatase, MKP-5."
RL J. Biol. Chem. 274.11949-11956(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20065165; PubMed=10597297;
RA Theodossiou A., Smith A., Gillieron C., Arkinstall S., Ashworth A.;
RT "WKP5, a new member of the MAP kinase phosphatase family, which
RT selectively dephosphorylates stress-activated kinases."
RL Oncogene 18:6981-6988(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain, Lung, and Testis;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.T., Wang J., Heish F.,
RA Diachenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Bork S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Richards S., McWay P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

```

```

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywicki M.I., Skalska U., Smalius D.E.,
RA Schneringer A., Schein J.E., Jones S.J.M., Maira W.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Involved in the inactivation of MAP kinases. Has a
CC specificity for the MAPK1/2/MAPK3/MAPK4 subfamily.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB026436; BA081668.1; -
DR EMBL: AF179212; AAD51857.1; -
DR EMBL: BC031405; AAH31405.1; -
DR HSSP: Q16828; MKP.
DR Genew: HGNC:3065; DUSP10.
DR GO: GO:0005737; Cytoplasm; TAS.
DR GO: GO:0005634; Cytosol; TAS.
DR GO: GO:0004721; P:protein phosphatase activity; TAS.
DR GO: GO:0007254; P:JNK cascade; TAS.
DR GO: GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR GO: GO:0006950; P:response to stress; TAS.
DR InterPro: IPR000340; DS phosphatase.
DR InterPro: IPR008343; MAPK phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR00387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS: PR01764; MAPKPHATASE.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANES 3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase: Nuclear protein.
FT DOMAIN 168 285 RHODANES.
FT PROTEIN-TYROSINE PHOSPHATASE.
FT DOKAIN 384 453 PHOSPHOCYSTINE INTERMEDIATE (BY
FT ACT_SITE 408 408 SIMILARITY).
SQ SEQUENCE 482 AA; 52642 MW; A8CB74ABF9498C04 CRC64;

```

Query Match 13.7%; Score 469; DB 1; Length 482;
Best Local Similarity 35.9%; Pred. No. 2.2e-21;
Matches 107; Conservative 65; Mismatches 96; Indels 30; Gaps 7;

```

QY 27 LIDSRPFVETNTSHILEATININCS-KLMKRILOQDKVITTELIO-HSAKHYVDIDCSOK 84
DB 173 VIIDCRPFMEVYKSHIQAVINADKISRRLQGGKITVLDISREGDSFKRIFSK 232
QY 85 VVVYVDSQD-VASLSSDCEFLVTLGLKLEKSPNSVHLIAGGFAEFSRCFPLCEGK----- 139
DB 233 IIVVDENETNSRVPQOPHIVLESIKRKGKPEVLVKGGLSFKKNHNLCDNSLQIG 292
QY 140 -----STLVPCISOPCLP-VANIGPTRIIPNLYIGQGVVANKELIQNGIGY 292
DB 233 CREVGAGSASSLLPPI-PTTPDIENALPTPIPLFLEINEDDADODLDTMQRITNGY 350
QY 188 VLNAS-----YTCKPDIPIESHFLRPVNDSCFCEKILFWLKDSDVDFIEKAKASNGCVLV 242

```

```

DB 351 VINVTHTLPLHYHEKGLF-----NYKRLPATDSNKQMLRQVFEBAFEBIAHOCGKGLI 406
QY 243 HCLAGISRSATTAIAIYMKRMDSLDEAVRFYKERTPTSPNPNFGQLIDYEKKIKN 300
DB 407 HCOAGVSRSAITVIAIYMKRMTMTDAYKFKVGRKPIISPINLFWGQLFEEDLNN 464

RESULT 7
DUSA_MOUSE STANDARD; PRT; 483 AA.
ID DUSA_MOUSE STANDARD; PRT; 483 AA.
AC Q9ES60; Q9CZY9;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 5) (MAP kinase phosphatase 5) (MKP-5).
DE DUSP10 OR MKP5.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20515587; PubMed=11060451;
RA Masuda K., Shima H., Kikuchi K., Matanabe Y., Matsuda Y.;
RT "Expression and comparative chromosomal mapping of MKP-5 gene DUSP10/Dusp10."
RL Cytogenet. Cell Genet. 90:71-74(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Araiwa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I., Saito T., Matzuda H.A., Ashburner M., Batalov S., Saito R., Kadoya K., Okazaki Y., Gojoubori T., Bono H., Kasukawa T., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Kochiwa H., Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Maehio T., Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barth G., Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Togo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlschki S., Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RT
CC -1- FUNCTION: Involved in the inactivation of MAP kinases. Has a specificity for the MAPK1/MAPK2/MAPK3/MAPK4 subfamily (by similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein + tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein + phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (by similarity).
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family. Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB037908; BAB17680.1; -.
DR EMBL; AK011995; BAB27966.1; -.
DR HSSP; Q16828; 1MKP.
DR MGD; MGI:1927070; Dusp10.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00450; RHOD; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS02026; RHODANSE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Nuclear protein.
FT DOMAIN 169 286 RHODANSE.
FT DOMAIN 386 436 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 409 PHOSPHOCYSTEINE INTERMEDIATE (BY SIMILARITY).
FT FT CONFLICT 48 48 A -> T (IN REF. 2).
FT FT CONFLICT 113 113 P -> S (IN REF. 2).
FT FT CONFLICT 386 386 R -> Q (IN REF. 2).
SQ SEQUENCE 483 AA; 52530 MW; 3551562355B46F89 CRC64;

Query Match 13.7%; Score 468; DB 1; Length 483;
Best Local Similarity 35.6%; Pred. No. 2,5e-21;
Matches 106; Conservative 67; Mismatches 95; Indels 30; Gaps 7;

QY 27 LIDSPPEVYNTSHLEAININCS-KLMKRLQDKVLITLIQ-HSAKHVDICQK 84
DB 174 VIIDCPFMEYNSKHIGAVHINCADKISRRLQCKITVLDISCREGDSFKRTFSKE 233
QY 85 VVYVYDSSQDVASLSDCETVLGLKLEKSFNSVHLLAGFAEFSCFPGLCGK----- 139
DB 234 IIVYDENTNEPSRVTPSPHLYVLESKREKGEPLVKGSLSEFKONHGNLCONSLOE 293
QY 140 -----STLVPTCISQCLP-VANIGPRLIPNLYLGGORDVNLKELIQNGTGY 187
DB 294 CREVGGAASASMLPQ--SVPTTPIENAEULPIPLFLFNGEDQAQDLDTMQLNIGY 351
QY 188 VLNAS-----YTCPKDPFIESHFLRVPNVDSFCEKILPWLKSVDFIEKASNGCVLV 242
DB 352 VINVTHTLPLHYHEKGLF-----NYKRLPATDSNKQMLRQVFEBAFEBIAHOCGKGLI 407
QY 243 HCLAGISRSATTAIAIYMKRMDSLDEAVRFYKERTPTSPNPNFGQLIDYEKKIKN 300
DB 407 HCOAGVSRSAITVIAIYMKRMTMTDAYKFKVGRKPIISPINLFWGQLFEEDLNN 465

RESULT 8
DUS6_MOUSE STANDARD; PRT; 381 AA.
ID DUS6_MOUSE STANDARD; PRT; 381 AA.
AC Q9DBB1; Q9D7L4;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase phosphatase 3) (MKP-3).
DE DUSP6 OR MKP3.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=liver, and Tongue;
RX MEDLINE=21085660; PubMed=11217851;

```


RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadoi K., Matsuda H.A., Ashburner T., Batcalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kiehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hormann M., Humé D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima M., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.,
RL "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loggellano N.A., Peters G.J., Abramson R.D., Millaby S.J.,
RA Bosak S.A., McMan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halys S.W.,
RA Villalón D.K., Muray D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R.W., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley A.C., Grimwood J., Schmutz J., Myers R.W.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002)
RL -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
family (By similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AK005062; BAB23766.1; -
DR EMBL; AK009131; BAB26093.1; -
DR EMBL; BC003869; AAH03869.1; -
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1914853; Dusp6.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DUSPc.1.
DR PRINTS; PR01764; MAPKPHPTASE.

DR SMART; SM00195; DUSPc.1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANESE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase.
KT DOMAIN 30 148 RHODANESE.
FT DOMAIN 206 381 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 293 293 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
FT CONFLICT 22 22 W -> G (IN REF. 1; BAB26093).
FT PROBLEM 34 34 L -> F (IN REF. 1; BAB26093).
SQ SEQUENCE 381 AA; 42407 MW; 7EALPBI54FPAAD2DA CRC64;
Query Match 13.6%; Score 466.5; DB 1; Length 381;
Best Local Similarity 33.7%; Pred. No. 2.2e-21;
Matches 112; Conservative 64; Mismatches 107; Indels 49; Gaps 8;
QY 19 LEQTEKVLIDSRPFVEVNTSHILRININCKMKRRLQDQKLTITLIGH-SAKHAY 77
DB 27 LELGNERLLMDCKRPODELSSHIESAINALVPIGMLRRLQKGNLPLRALTFRCEDRDHF 86
QY 78 DIDC-SQKVVVYVYDQSSQD-VASLSDCFLTVLLGLKLEKSFNSVHLIAGGFAPFRCPGL 135
DB 87 TRRCGTDTVVLVDENSSDMNENNGSGSVGLLKLKDKGCAFTYEGGFSKFGAEFALH 146
QY 136 CEKSKTVTPVC-ISQCLPVPANIG----- 158
DB 147 CE--TWLIDGSCSSSPPLPLVLTGGLRISDSSTDIESDLDRDPNSATDSDSPLSNSQP 204
QY 159 --PTRLIPNLYAGCORVDYANKELIQNGIGVYANASYCPKP-DLPESHPLRPVNPNSF 215
DB 205 SFVELLPFLYLGCADSTNLVDLEFGIKTLINVTNPLNLPENNGEFTKQIPLSDHW 264
QY 216 CEKILPWLDSVDFIEKAVASNGCVLVHCLAGISRSATIAIYIMKRMDSIDEAYRPFK 275
DB 265 SQMLISQFFPFAISFIDEARGKNCVYLHCLAGISRSVTVYVLMQNLNMDAYDYLK 324
QY 276 EKRPSTPNPNFLGOLDYEXKIKKQYTGASGP 307
DB 325 MKKSNISPNFNPFGQLDPERTL---GLSSP 352
RESULT 9
DUS6_RAT
ID DUS6_RAT STANDARD; PRT; 381 AA.
AC 064336;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
DE phosphatase 3) (MKP-3).
DE DUSP6 OR MKP3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Neuron;
RX MEDLINE=96224012; PubMed=8626780;
RA Muda M., Boschart U., Dickson R., Martinou J.C., Martinou I.,
RA Camps M., Schlegel W., Arkinsfall S.;
RT "MKP-3, a novel cytosolic protein-tyrosine phosphatase that
RT exemplifies a new class of mitogen-activated protein kinase
RT phosphatase.";
RL J. Biol. Chem. 271:4319-4326(1996).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96216487; PubMed=8631996;
RA Mourey R.J., Vega Q.C., Campbell J.S., Wenderoth M.P., Hauschka S.D.,

RA Krebe E.G., Dixon J.E.;
 RT "A novel cytoplasmic dual specificity protein tyrosine phosphatase
 RL J. Biol. Chem. 271:3795-3802(1996).
 CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
 CC family. Implied in muscle and neuronal differentiation.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL: X94185; CA63895.1; -;
 CC EMBL: U42627; AAB06202.1; -;
 CC HSSP: Q16828; 1MKP.
 CC InterPro: IPR000340; DS_phosphatase.
 CC InterPro: IPR008343; MAPK_phosph.
 CC InterPro: IPR001763; Rhodanese-like.
 CC InterPro: IPR000387; TYR_phosphatase.
 CC Pfam: PF00782; DSpC; 1.
 CC Pfam: PF00581; Rhodanese; 1.
 CC PRINTS: PR01764; MAPKPHPTASE.
 CC SMART: SM00195; DSpC; 1.
 CC SMART: SM00450; RHOD; 1.
 CC PROSITE: PS50206; RHODANES_3; 1.
 CC PROSITE: PS50383; TYR_PHOSPHATASE_1; FALSE_NEG.
 CC PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 CC PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 CC HydroLase.
 CC DOMAIN 30 148 RHODANES.
 CC PROTEIN-TYROSINE PHOSPHATASE.
 CC ACT_SITE 293 293 PHOSPHOCYSTEINE INTERMEDIATE (BY
 CC SIMILARITY).
 CC SEQUENCE 381 AA; 42318 MW; C511E0CB68F2888 CRC64;
 CC -----
 CC Query Match 13.6%; Score 465.5; DB 1; Length 381;
 CC Best Local Similarity 33.7%; Pred. No. 2.6e-21;
 CC Matches 112; Conservative 64; Mismatches 107; Indels 49; Gaps 8;

RESULT 10
 ID DUS4_HUMAN STANDARD; PRT; 394 AA.
 AC Q13115; Q13524;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
 DE phosphatase-2) (MKP-2) (Dual specificity protein phosphatase hwh2).
 GN DUSP4 OR MKP2 OR VH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95221370; PubMed=7535768;
 RA Guan K.-L., Butch E.;
 RT "Isolation and characterization of a novel dual specific phosphatase,
 RT HWH2, which selectively dephosphorylates the mitogen-activated
 RT protein kinase.";
 RL J. Biol. Chem. 270:7197-7203 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96198119; PubMed=8626452;
 RA Chu Y., Solaki P.A., Khosravi-Far R., Der C.J., Kelly K.;
 RT "The mitogen-activated protein kinase phosphatases PAC1, MKP-1, and
 RT MKP-2 have unique substrate specificities and reduced activity in vitro
 RT toward the ERK2 severnaker mutation.";
 RL J. Biol. Chem. 271:6497-6501 (1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Skin, and Uterus;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diachenko L., Marusha K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleorn M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Tohyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abrahams R.D., Mulhally S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Regulates mitogenic signal transduction by
 CC dephosphorylating both Thr and Tyr residues on MAP kinases
 CC ERK1 and ERK2.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
-----
DR EMBL: U21108; AAA85119.1; -.
DR EMBL: U48807; AAC50452.1; -.
DR EMBL: BC002671; AAH02671.1; -.
DR EMBL: BC014565; AAH14565.1; -.
DR HSSP: Q16828; IMKP.
DR Genew: HGNC:3070; DUSP4.
DR MIM: 602747; -.
DR GO: GO:0005634; C:nucleus; TAS.
DR GO: GO:0008330; F:protein tyrosine/threonine phosphatase acti. .; TAS.
DR GO: GO:0000165; P:MAPKK cascade; TAS.
DR GO: GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PR00782; DSPC; 1.
DR Pfam: PR00581; Rhodanese; 1.
DR PRINTS: PR01764; MAPKPHPTASE.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANES_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR HydroLase; Nuclear protein.
KW DOMAIN 41 159 RHODANES.
FT DOMAIN 197 394 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 280 280 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY)
SQ CONFLICT 111 111 R -> G (IN REF. 2).
SEQUENCE 394 AA; 42953 MW; 0603971759B6952E CRC64;

Query Match 13.0%; Score 445; DB 1; Length 394;
Best Local Similarity 31.4%; Pred. No. 4.8e-20;
Matches 120; Conservative 75; Mismatches 141; Indels 46; Gaps 14;

QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCKLMKRLQODKVLITELI--QHSAKHK 76
DB 39 LPSG-GKCLLDRCRFLAHAGYILGVSIVNRCNTIVRR-AKGSVSEQLIPAEVEVRAR 96
QY 77 VIDDSOKVVYDQSSQVVASLSDCFILVLGKLEKSF--NSVHLAGGPAEFSRCPFG 134
DB 97 LRSGLYSAYIVYDERSPPAESLRDSTVSLVQALRRNAERTDCLKGGYERFSSEYPE 156
QY 135 LCEGKSTL-----VPTGISQP-----C-LPVANI-GPTRIIPNTYLGQRVNLKEL 179
DB 157 FCSKTKALAIIPPVPSPATELELDGSCSGPTLHDGQFVEILLPFIYLSAHTAARDM 216
QY 180 IQQNGIGVLAASYCEKPDTPESHFLRVNDVSCEKILPWLDKSVDFTEKAKANGC 239
DB 217 LDAIGITALLNVSSDCPN-HFEGHVQYKCIPEVDNHADISSWMEAEIYIDAVKDCGR 275
QY 240 VLVEGLAGISRSATIAIYIKRMDMSIDEAYRVKKEKPTISNENFLQQLDYEKIK 299
DB 276 VLIVHQAQGISRSATITCLAYLMKKRVIRBEAFVKKPRSTISNFSFMQQLQFESQVL 335
QY 300 NOTGASGPKSLKLLHKEKPEVPVAVSEGGQSETELPSCADSAESEAAGRPVYPAS 359
DB 336 ATSCAAEAS-----PSGFL-----REKGTATAT-----TSQVFSEFPV-SVG 373
QY 360 VPSVPSQPSLSDSPVLQALS 381
DB 374 VHSAPSSLPYL--HSPITTSPS 393
```

```
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DB Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
DB (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
DB phosphatase-2) (MKP-2).
DB DUSP4 OR MKP2.
GN Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=20379359; Pubmed=10918612;
RA Fu S.-L., Wana A., Vogt P.K.,
RT "Identification and characterization of genes upregulated in cells
RT transformed by v-Jun."
RL Oncogene 19:3537-3545(2000).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases ERK1
CC and ERK2 (By similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
-----
DR EMBL: AF167296; AAD4656.1; -.
DR HSSP: Q16828; IMKP.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PR00782; DSPC; 1.
DR Pfam: PR00581; Rhodanese; 1.
DR PRINTS: PR01764; MAPKPHPTASE.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANES_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW HydroLase; Nuclear protein.
FT DOMAIN 25 143 RHODANES.
FT DOMAIN 178 375 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 261 261 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY)
SQ SEQUENCE 375 AA; 41052 MW; 179290DC2BEEEF1 CRC64;

Query Match 13.0%; Score 444; DB 1; Length 375;
Best Local Similarity 31.1%; Pred. No. 5.1e-20;
Matches 123; Conservative 73; Mismatches 152; Indels 48; Gaps 13;

QY 4 EMIGQIVTERVALLSGETKVLIDSRPVEYNTSHILEAININCKLMKRLQODKV 63
DB 9 EMEGSAL--RLVGRREBSGRCCLLDRCRFLAHAGYILGVSIVNRCNTIVRR-AKGAV 65
QY 64 LITELI--QHSARKKVIDDSOKVVYDQSSQVVASLSDCFILVLGKLEKSF--NSVH 119
DB 66 SLBQILPAEGEVRRARLAKGYTAVVLYDERSPPAESLRDSTVSLVQALRRALRRMADADR 125
```

QY 120 LLAGFAEBSRCFPGICEGKSTLVPTICISOP-----C-LPVANI--GPTRIILPN 165
DB 126 LLAGYERFSEPEPCAKTKTL--SISPSASIESIDLQSSCGTPPLHDQGPVEILLP 183
QY 166 LYLGGCRDVLNKLKIQNGIGVYLNASYTPCKPDPIESHFLRPVNDSPCEKILPWLDK 225
DB 184 LYLGSAYHARRMDALGITALLNVSSDCPN-HFEHGYKCIPIEDNHKALISSWFM 242
QY 226 SVDFIEKAKASNCVLVHCLAGISRSATIAIYIMKMDNSLDEAYRFVKEKPTISPNF 285
DB 243 AIEIYDSVEKCCGRVLVHCOAGISRSATICTAYIMMKRVKLEKAEFEVQORSIISPNR 302
QY 286 NFIQGLLDYKRTIKNGTASGSPSKLKLHLEKNEVPVAVSEGGQSETPPLSPCADSA 345
DB 303 SFNGQLQFESQVLTATSCAVEAAS-----PSGTLRERKATSTP----- 341
QY 346 TSEAGQRPVHPASVPSVPSQPSLDESPVLQALS 381
DB 342 TSQFVFSFPV-SVGVHATPSSLPYL--HSPITTS 374

RESULT 12
DUS4_RAT STANDARD; PRT; 395 AA.
AC 062757;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (BC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase phosphatase-2) (MKP-2).
GN DUSP4 OR MKP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Phenochromocytoma;
RX MEDLINE=95301550; PubMed=7782322;
RA Maira-Press A., Rim C.S., Yao H., Roberson M.S., Stork P.J.S.;
RT "A novel mitogen-activated protein kinase phosphatase. Structure, expression, and regulation.";
RT J. Biol. Chem. 270:14587-14596(1995).
CC -1- FUNCTION: Regulates mitogenic signal transduction by dephosphorylating both Thr and Tyr residues on MAP kinases ERK1 and ERK2 (by similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein + phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed at moderate levels in nearly all tissues and cells including brain, spleen, and testes with the higher expression in the heart and lung and lower expression in skeletal muscle and kidney. Undetectable in liver. Expressed in many areas of the brain with very strong expression in the hippocampus, piriform cortex, and the suprachiasmatic nucleus.
CC -1- INDUCTION: By mitogens and by stress.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; U23438; AAC52493.1; -.

DR HSSP; Q16828; MKP.
DR Interpro; IPR000340; DS_phosphatase.
DR Interpro; IPR008343; MAPK_phosph.
DR Interpro; IPR001763; Rhodanese-like.
DR Interpro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00782; DSCP; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHRTASE.
DR SMART; SM00195; DSCP; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydroxylase; Nuclear protein.
FT DOMAIN 42 160 RHODANES.
FT DOMAIN 198 395 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 281 281 PHOSPHOCYSTEINE INTERMEDIATE (BY SIMILARITY).
SQ SEQUENCE 395 AA; 43187 MW; A90EFPD378A050FD CRC64;
Query Match 12.9%; Score 439.5; DB 1; Length 395;
Best Local Similarity 30.5%; Pred. No. 1e-19;
Matches 118; Conservative 71; Mismatches 147; Indels 51; Gaps 12;
QY 16 VALLESGETKLVLLIDRPFEVNTSHILEAININGSKMKRLQDQKVLITELI--QHSA 73
DB 38 LGLLSGG--KCLLLDQPLASAGYIRGSVNVRCITVRR--AKGSVLEQILPAEEV 94
QY 74 KHKVDIDCSQKVVVYVYDOSQDVASLSSDFLVTLGLKLEKSF--NSVHLLAGFAEBSRC 131
DB 95 RARLRSGLYSANIVYVDERSPRAESLREDSTVLVQALRRNERTDICLKGYEFSSE 154
QY 132 FPGLCGKSTLVPTICISOPCLPVANI-----GPTRIILPNLYGQQRDV 174
DB 155 YEFSCSKTKAL--AAPRPVPSTNIESLIDGSSCGTPPLHDQGPVEILPFLVLSAYHA 212
QY 175 LNKELIQNGIGVYLNASTYCKRPDPIESHFLRPVNDSPCEKILPWLDKSDPIEKAK 234
DB 213 ARDDMDALGITALLNVSSDCPN-HFEHGYKCIPIEDNHKADISSWMEALEYIDAVK 271
QY 235 ASNGCVLVHCLAGISRSATIAIYIMKMDNSLDEAYRFVKEKRPISPNFPLQGLDY 294
DB 272 DCRGRVLVHCOAGISRSATICTAYIMMKRVLEKAEFEVQORSIISNFSMGQLLP 331
QY 295 EKKIKNGTASGSPSKLKLHLEKNEVPVAVSEGGQSETPPLSPCADSATSEAGQRP 354
DB 332 ESQVLTSCAVEAAS-----PSGF--LRERKATPTP-----TSQFVFSFP 370
QY 355 VHPASVPSVPSVPSQPSLDESPVLQALS 381
DB 371 V-SVGVHAPPSNLPYL--HSPITTS 394

RESULT 13
DUS1_MOUSE STANDARD; PRT; 367 AA.
AC P28563;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (BC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase ERP).
GN DUSP1 OR PTPN10 OR MKP1 OR 3CH134 OR PTPN16.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=92158357; PubMed=1741163;

RA Charles H., Ahler A.S., Lau L.F.;
RT "cDNA sequence of a growth factor-inducible immediate early gene and
RL characterization of its encoded protein."; Oncogene 7:187-190(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=93360956; PubMed=8355678;
RA Noguchi T., Metz R., Chen L., Mattei M.-G., Carrasco D., Bravo R.;
RT "Structure, mapping, and expression of erp, a growth factor-inducible
RT gene encoding a nontransmembrane protein tyrosine phosphatase, and
RL effect of ERP on cell growth."; Mol. Cell. Biol. 13:5195-5205(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shmamen C.M., Schuler G.D.,
RA Alschul S.F., Zeeberg B., Butler K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hasteh F.,
RA Ditchenko L., Marinina K., Farmer A.C., Rubin G.M., Hong L.,
RA Stepienon M.J., Soares W.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshlyuk S., Carninci P., Franke C.,
RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halysk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey U., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulfard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodrigoquez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.B.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=94037096; PubMed=8221888;
RA Sun H., Charles C.H., Lau L.F., Tonks N.K.;
RT "MKR-1 (3CH13), an immediate early gene product, is a dual
RL specifically phosphatase that dephosphorylates MAP kinase in vivo."; Cell 75:487-493(1993).
CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
CC kinase ERK2 on both Thr-183 and Tyr-185.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- INDUCTION: By growth factors.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanase domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on it's
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL, X61940; CAA43944.1; -
DR EMBL, S64851; AAB27882.1; -
DR EMBL, BC006967; AAH06967.1; --
DR PIR, A54681; S24411.
DR HSSP, O16828; IMKP.
DR MGJ, MGI:105120; Dusp1.
DR InterPro, IPR000340; DS_phosphatase.
DR InterPro, IPR008343; MAP_kinophos.
DR InterPro, IPR001763; Rhodanese-like.
DR InterPro, IPR000387; TYR_phosphatase.

DR	pfam; PF00782; DSPC; 1.
DR	pfam; PF00581; Rhodanese; 1.
DR	PRINTS; PR01764; MAPKPHATASE.
DR	SMART; SMO0195; DSPc; 1.
DR	SMART; SMO0450; RHOD; 1.
DR	PROSITE; PS50206; RHODANESE_3; 1.
DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR	PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR	PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW	Hydrolase; Cell cycle.
FT	DOMAIN 20 137
FT	DNAIN 175 367
FT	ACT SITE 258 258
FT	MUTAGEN 258 258
SQ	SEQUENCE 367 AA; 39369 MW; 50B5FP0FEBBD19AB CRC64;
Query Match 12.7%; Score 433; DB 1; Length 367;	
Best Local Similarity 31.6%; Pred. No. 2.3e-19;	
Matches 115; Conservative 54; Mismatches 151; Indels 44; Gaps 8	
QY	15 LVALLSGTEKVLLIDSRPFVEVNTSHILFAININCSTKMRRRLQQDQKYLITELIQHSAK 74
DB	13 LRALLREGAACILLDCRSFFAPPAAGHAGSVNRFSTIVARRKAGMGLHEIIVNAELR 72
QY	75 HKVDIDSQRVVYDQSQDVASLSSDCELTVLIGKL--EKSFNSVHLLAGFAEFSRCF 132
DB	73 GRLLTAGVAHVAVLIDERSASIDGAKRDGTLLAAGALCREARSTGVFLQGYEAFASAC 132
QY	133 PGLEEGESTLVPTGISQP-----CLTPVANI-----GPRIIPNLVGQQRDL 175
DB	133 PELCSKOST--PTGLSLPLSTVSVPDSABSGSSCSTPIYDGGPVELLIPLYLSGAAYAS 190
QY	176 NKEELIQNGIGYVLAASYTCPEKDPFIPEBSHFARPVNDSCFEKILPWLDKSYDFIEKAKA 235
DB	191 RKMULDMLGLVALINVANSNCN-HPEGHYQXKSIFVEDNHADISSWNEAIEDFDISTKD 249
QY	236 SNGCVLVHCLAGIRSRTIALAIYIMKRMDMLDEAYRVFKERKPRISPNFNFGQLDYE 295
DB	250 AGGRVFHYCQGIGIRSATICLAYIMRNRYKLDEAFEFVKQRSITISPSFSWGQLQPE 309
QY	296 KIKINQTGASGPSKTLKILHEKENEPVPAVSEGQSKETPLSPCADSATSEAGORPV 355
DB	310 SQV-----LAPHGS-----AEKGSPMAVALDRGTSTTYVNFPVS-----IPV 347
QY	356 HPAS 359
DB	348 HPTN 351
RESULT 14	
ID	DUST_RAT STANDARD; PRT; 367 AA.
AC	O64623;
DT	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE	(MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase CL100).
GN	DUSP1 OR CL100.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OK	NCBI_TaxID=10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Sprague-Dawley; TISSUE=Lung;
RA	Muda W., Schlegel W., Atkinson S.;
RL	Submitted (JAN-1995) to the EMBL/Genbank/DBJ databases.
CC	-1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
CC	kinase ERK2 on both Thr-183 and Tyr-185.
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC	tyrosine + phosphate.

CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
phosphate.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X64004; CAA58828.1; -.
DR PIR: S52265; S52265.
DR HSSP: Q16828; 1MKP.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSEC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR SMART: SM00195; DSEC; 1.
DR SMART: SM00450; RHOD; 1.
DR SMART: SM00206; RHODANES; 3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydrolyase; Cell cycle.
FT DOMAIN 20 137 RHODANES.
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE (BY
SIMILARITY).
SQ SEQUENCE 367 AA; 39541 MW; 5112ADF290499139 CRC64;
Query Match 12.4%; Score 423; DB 1; Length 367;
Best Local Similarity 31.3%; Pred. No. 9.4e-19;
Matches 115; Conservative 54; Mismatches 148; Indels 50; Gaps 9;
QY 15 LVALLSGTEKVLIDSRPEVNTSHLEAININCSKIMKRLQ-----QDKVLTLEIL 69
DB 13 LRLALLERAAQCILLDRSFFAFNAGIVSVVVRSTIYRRKAKMGLEHIVPTLEL- 71
QY 70 QHSARKHKVDIDCSQKVVVYDQSSQDVASLSDCFIVLGLKTL- EKSPNSVHLLAGSFAE 127
DB 72 ----RGRLLAGAYHAVVLDERSAALDGAKRGGTLAAGALCREARSTQVFLGGGYEA 127
QY 128 FSRCPFGLECGK-----STVPTCISQPCLPVANI-----GPTRLPMLYIGCOR 172
DB 128 FSASCEBELSCSKOSTPMGLSLPLSTVSDASQSCSSCPPLVDQGPVELSLFLYGSAY 187
QY 173 DVLNKEILQONGNGVYLNAASYTPKPDFIPESHFLRPVVDSCCKILPMLDSVDFIEK 232
DB 188 HASRKMDLADALGITALINVSANCPN-HFBSHYQYKSIPEVDNKADISSFENAFIDFS 246
QY 233 AKASNCVVLVHCLAGISRSATIAIYIMKMDLSLEAYRFEVKEKPTISPNFNFLGOLL 292
DB 247 IKDAGRVFHCQAGISRSATICLAVLMRTNRKYLDEAFEFVQGRSIIISPNFSGOLL 306
QY 293 DYKATIKNQTGASGPKSKLTLHLEKNEPVPVVSBCQKSETPPLSPCCADATSBAAG 352
DB 307 QFESSQV-----LAFHCS-----AEGSPAMAVLDKRGSTTVTFNFPVS----- 344
QY 353 RPYHPAS 359
DB 345 IPVHPTN 351

AC P28562;
DT 01-FEB-1992 (Rel. 24, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
DE CL100) (Dual specificity protein phosphatase hVH1).
GN DUSP1 OR PRPNI1 OR MKP1 OR CL100 OR VH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ON NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Forebrain;
RX MEDLINE=93024952; PubMed=1406996;
RA Keyes S.M., Emalle E.A.;
RT "Oxidative stress and heat shock induce a human gene encoding a
RL protein-tyrosine phosphatase.";
RL Nature 359:644-647 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Pange C.,
RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mollay S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
RA Villalón D.K., Munzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Ketterman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schercher A., Schein J.E., Jones S.J.M., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
CC kinase ERK2 on both Thr-183 and Tyr-185.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- INDUCTION: By oxidative stress and heat shock.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68277; CAA48338.1; -.
DR EMBL: BC022463; AA422463.1; -.
DR PIR: S29090; S29090.
DR HSSP: Q16828; 1MKP.
DR Genew; HGNC:3064; DUSP1.
DR MIM; 600714; -.
DR GO; GO:0004726; P:non-membrane spanning protein tyrosine phos. . .; TAS.
DR GO; GO:0006979; P:response to oxidative stress; TAS.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.

```
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS0206; RHODANES 3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE 1; 1.
DR PROSITE; PS0056; TYR_PHOSPHATASE 2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase; Cell cycle.
FT DOMAIN 20 137 RHODANES.
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 367 AA; 39297 MW; 11BD1D39A9FCD51P CRC64;

Query Match 12.3%; Score 420; DB 1; Length 367;
Best Local Similarity 31.2%; Pred. No. 1.4e-18;
Matches 113; Conservative 56; Mismatches 159; Indels 34; Gaps 7;

QY 15 LVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKMKRRLOQDKVLTITELIQHSK 74
DB 13 LRALGERAAQCLLDRCSPFAFNAGHAGSVNVAFTIVRRRAKAMGLEHIVNMLR 72
QY 75 HKVDIDCGQKVVVDQSQDVASLSDCFLTVLIGKL--EKSFNSVHLLAGPAEFGRCF 132
DB 73 GRLLAGAYHAYVLDERSAALDGAARDGTLLALAGALCREARAQVFFLKGGYENFASAC 132
QY 133 PGLCEGK-----STLVPTCISQPLPVANI-----GPTRIIPNLVLGCRDVLNK 177
DB 133 PELCSKQSTPMGLSLPLSTSVPSDAESGSCSSTPLYDQGVLEILPFLYLGSAYHASRK 192
QY 178 ELIQNGIGYVLNASTYCPKPDPIPESHFLKVPVNDSCFKILFWLDSVDFIEKAKASN 237
DB 193 DMLDALGITALINVSANCPN-HFEGHYQKSIPEVDNHKADISSWFNEAIDFIDSINKAG 251
QY 238 GCVLVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPTISPNFNFLOGLLDYEKK 297
DB 252 GRVFNHCQAGISRSKTICTLAITMKTNRVYKDBAEFVQKRSIISPNSFMQQLQFESQ 311
QY 298 IKNQTGASGPKSKLTLHLKKNPEVPVAVSEGGQKSET---PLSPPCADSATSEAAQQR 353
DB 312 V-----LAPHCS-----AEAGSPAMAVLDRGTSTTVTFNFPVSIIVHSTNSALSTYQS 359
QY 354 PV 355
DB 360 PI 361
```

Search completed: June 21, 2004, 13:21:07
Job time : 16.4416 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 20:53:18 ; Search time 4983.03 Seconds
(without alignments)
3985.196 Million cell updates/sec

Title: US-10-029-345A-109
Perfect score: 3418
Sequence: 1 MAHEMGTQVTRVALLE.....LGRVSGSSRSGSMETIEVS 665

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=/cgr2/1/USPTO.spool/US10029345/rnat 21062004.164146.29039/app_query.fasta.1.1294
-DB=EST -CPMT=fastap -SUFFIX=ret -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345.0CCN.1.1.3931.0rnat 21062004.164146.29039 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_g88_hum:*
18: em_g88_hiv:*
19: em_g88_pln:*
20: em_g88_vrt:*
21: em_g88_fun:*
22: em_g88_man:*
23: em_g88_mus:*
24: em_g88_pro:*
25: em_g88_rtd:*
26: em_g88_phg:*
27: em_g88_vrl:*
28: gb_g88l:*

29: gb_g882:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1553	45.4	1060	12	BM546940
2	1412	41.3	881	13	BQ30140 AGENCOURT
3	1397.5	40.9	883	13	BQ945892 AGENCOURT
4	1375.5	40.2	974	13	BQ951695 AGENCOURT
5	1375	40.2	859	13	BQ933499 AGENCOURT
6	1369.5	40.1	898	13	BQ72165 AGENCOURT
7	1360	39.8	903	12	BQ482429 AGENCOURT
8	1327	38.8	2207	11	BC038231 Homo sapi
9	1306	38.2	862	13	BQ770036
10	1216	35.6	769	13	BU704078
11	1180	34.5	836	12	BI821804
12	1148	33.6	715	14	CF727177
13	1102.5	32.3	920	10	BE897795
14	1086.5	31.8	682	12	BI16954
15	1081	31.6	656	14	CF532917
16	1076	31.5	639	14	CF742387
17	1075	31.5	891	13	BQ670989
18	1073.5	31.4	946	12	BI657528
19	1071.5	31.3	660	12	BI917706
20	1065	31.2	983	12	BG165188
21	1059.5	31.0	790	12	BI661614
22	1047	30.6	663	13	BK479029
23	1029.5	30.1	752	13	BQ432082
24	1006	29.4	650	10	BF472046
25	979	28.6	621	9	AV703072
26	978.5	28.6	924	14	CA474739
27	972.5	28.5	743	14	CF727802
28	970	28.4	693	14	AV701628
29	964.5	28.2	739	28	AZ850283
30	964.5	27.4	602	10	AM847425
31	931	27.2	595	10	AM847426
32	925	27.1	631	28	BH039241
33	911	26.7	1067	10	BF135687
34	891	26.1	3325	11	AK035652
35	872	25.5	792	13	BU946569
36	839.5	24.6	586	14	CF104291
37	828	24.2	723	13	BU708175
38	823.5	24.1	516	29	AY413655
39	804.5	23.5	516	29	AY413657
40	772.5	22.6	683	12	BM942528
41	770	22.5	619	29	CG583869
42	769	22.5	579	10	BF582366
43	768.5	22.5	594	14	CF727694
44	762	22.3	512	10	BF815601
45	749.5	21.9	538	10	AW909957

ALIGNMENTS

RESULT 1
BM546940
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

BM546940
AGENCOURT_6491295 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5723711
5', mRNA sequence.
BM546940
BM546940.1 GI:18780310
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1060)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
[found through the I.M.A.G.E. Consortium/LLNL at:](http://image.llnl.gov)
<http://image.llnl.gov>
 Plate: LHM12711 row: m column: 24
 High quality sequence stop: 737.
 Location/Qualifiers
 1. 1060

FEATURES
 source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone IMAGE:5723711"
 /lab_host="DH10B"
 /clone_11b="NIH MGC 125"
 /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
 Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
 of three ovaries, from females ranging in age from 38 to
 49 yo. Library is oligo-dT primed and directionally cloned
 size 2.1 kb, insert size range 1-3.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 036."

ORIGIN

Alignment Scores:

Pred. No.: 2,95e-125 Length: 1060
 Score: 1553.00 Matches: 317
 Percent Similarity: 93.91% Conservative: 7
 Best Local Similarity: 91.88% Mismatches: 15
 Query Match: 45.44% Indels: 7
 Gaps: 2

US-10-029-345a-109 (1-665) x BM546940 (1-1060)

QY 163 LeuProAsnLeuTyrlleuGlyCysGlnAaSPValleuAsnlyGluLeuIleGlnGln 182
 Db 6 ATTCCCAATCTTATCTGTGCTGCAGCAGATGCTCCCAAGAGAGCTGATGACAG 65
 QY 183 AaGlyIleGlyTyrlleuAsnAaSerTyrlleuProAspPheIlePro 202
 Db 66 AATGGGATTGTTATGTTAAATCCCAAGAACCTGTCMAAGCCTGATTCCTCC 125
 QY 203 GluSerHisPheLeuAaValProValAsnAspSerPheCysGluTyrlleuProT 222
 Db 126 GAGCTCTCATTTCTCGTGTGCTGATGACAGCTTTGTGAGAAATTTTGGCCGTG 185
 QY 223 LeuAspLysSerValAspPheIleGluTyrlleuAaValAsnSerGlyCysValleu 242
 Db 186 TTGGCAATATGATGATTCATTCAGAAAGCAAGCCCTCCATGATGATGTTAGTG 245
 QY 243 HisCysLeuAaGlyIleSerAspSerAlaThrIleAlaIleAlaTyrlleuTyrlleu 262
 Db 246 CACTGTTTACGTGGAGATCTCCCGCTCCGACCATGCTATGCCATCATGAGAGG 305
 QY 263 MetAspSerSerLeuAspGluAaTyrlleuTyrlleuAaValAsnSerGlyCysValleu 282
 Db 306 ATGACACATGCTTTAGATGAGAGCTTCAGATTTGTGAGAAAGAAAGACCTACTATATCT 365
 QY 283 ProAsnPheAsnPheLeuGlyIleLeuLeuAspTyrlleuTyrlleuAaValAsnGln 302
 Db 366 CCAAACTTCAATTTCTGGCCAACTCTGACATGAGAGAGATTTAAGAACCAAGACT 425
 QY 303 GlyAlaSerGlyProLysSerLysLeuLysLeuHisIleuGluTyrlleuProAsnGluPro 322

Db 426 GCAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGCACCTGAGAAAGCCAAATGAACCT 485
 QY 323 ValProAlaValSerGluGlyIleGlyIleLysSerGluThrProLysSerProProVal 342
 Db 486 GTCCCTGCTGTCTCAGAGGTGACAGAAAGCCAGACGCCCTCATCTCAGCTGTGCC 545
 QY 343 AspSerAlaThrSerGluAlaAlaGlyIleAspProValHisProAlaSerValProSer 362
 Db 546 GACTGTGCTACTCAGAGCAGCAGCAAGAGCCCTGATCCCGCC-----AGC 596
 QY 363 ValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeuSerGly 382
 Db 597 GTACCCAGCGTACAGCCGTGCTGTGATGAGACAGCCCGCTGTGACAGCGCTCATGGG 656
 QY 383 LeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysAspSerPheSerLeu 402
 Db 657 CTGCACCTGTCCGACAGCAGCTGGAAGACAGCAATAGCTCAAGCGTTCTCTCTCG 716
 QY 403 AspIleLysSerValSerTyrlleuSerAlaSerMetAlaAlaSerLeuHisGlyPheSerSer 422
 Db 717 GATATCAAAATCAAGTTCAATTCAGCCAGCAGCAGCATCTTTCATAGGCTTCTCTCA 776
 QY 423 SerGluAspAlaLeuGluTyrlleuTyrlleuProSerThrThrLeuAspGlyThrAsnLysLeu 442
 Db 777 TCAGAGATGCTTTGGAATCACTCAAACTTCACTCACTGAGATGGACCAACAACTCA 836
 QY 443 CysGlnPheSerProValGlnLeuSerGluGlnThrProGluThrProAspLys 462
 Db 837 TGCCATTTCTCCCGTCCGATCAGAACTATCGAGAGAGACTCCGAAACAGATCT-GATTAAG 895
 QY 463 GluGluAlaSerIleProLysLysLeuGlnThrAlaArgPro-SerAspSerGlnSerIle 482
 Db 896 GAGAGAGCCAGCATCCCGCAGAGCTGACAGCCCGGCTTCAACAGCAGCAAGAGA 955
 QY 482 sArg-LeuHisSerValArgThrSerSer---GlyThrAlaGlnArgSerLeuLeu 501
 Db 956 GCGATTCATTCATTCAGTATGAAACCCAGCAGCAGTGGCAGCCGCCAGAGATGCCCTTT 1015
 QY 501 erProLeuHis 504
 Db 1016 TTATCTCTCCAC 1026

RESULT 2
 BQ30140
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 881)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE
 JOURNAL
 COMMENT
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ruben Laboratory
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
[found through the I.M.A.G.E. Consortium/LLNL at:](http://image.llnl.gov)
<http://image.llnl.gov>
 Plate: LHM2648 row: h column: 21
 High quality sequence stop: 676.
 Location/Qualifiers
 1. 881

FEATURES

source
 /organism="Homo sapiens"
 /mol_type="mRNA"

/db_xref="taxon:9606"
 /clone="IMAGE:6460964"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH MGC 101"
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 4,98e-113 Length: 881
 Score: 1412.00 Matches: 285
 Percent Similarity: 98.62% Conservative: 0
 Best Local Similarity: 98.62% Mismatches: 1
 Query Match: 41.31% Indels: 3
 Gaps: 0

US-10-029-345A-109 (1-665) x BQ930140 (1-881)

QY 266 SerLeuApGluAlaTyrArgPheValIyGluLysArgProThrIleSerProAspPhe 285
 DB 3 TCTTTAGATGAGCTTACAGATTGTGAAAGAAAAAGACCTTATATCTCCAAACTTC 62
 QY 286 AenPheUenGlyGlnLeuLeuAapTyrgLylLysIleLysAenGlnThrgLylAser 305
 DB 63 AATTTCCTGGCCCAACTCTGACATAGAGAGAGATTAAGAACCAAGACTGGAGATCA 122
 QY 306 GlyProLysSerLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLys 325
 DB 123 GGGGCAAGAGCAAACTCAAGCTGCTCACTGGAGAGCAATGAACTGCTCCCTGCT 182
 QY 326 ValSerGluGlyGlnLysSerGluThrProLysSerProProCysAlaAspSerAla 345
 DB 183 GTCTCAGAGGGGTGACGAAAGAGGAGACCCCTCAGTCACCCCTGTGCGACTGCT 242
 QY 346 ThrSerGluAlaIaGlyGlnArgProValHisProAlaSerValProSerValProSer 365
 DB 243 ACCTCAGAGGAGAGGAGCAAGGCCCCGTGATCCGCCAGCGCCACGCTGCCAGC 302
 QY 366 ValGlnProSerLeuLeuGluAapSerProLeuValGlnAlaLeuSerGlyLeuHisLeu 385
 DB 303 GTCCAGCCGTCGCTTAAAGAGACAGCCCGCTGTACAGCGCTCACTGGGCTGCACTG 362
 QY 386 SerAlaAspArgLeuGluAapSerAenLysLeuLysArgSerPheSerLeuAapLys 405
 DB 363 TCCGCAAGACAGCTGGAAGACAGCAATTAAGCTCAAGCGTCTCTCTCGATATCAAA 422
 QY 406 SerValSerLysSerAlaSerMetAlaAlaSerLeuHisGlyPheSerSerSerGluAap 425
 DB 423 TCAGTTCAATTCAGCCAGATGCGAGCAATCTTACATGCTCTCTCATCAAGAGAT 482
 QY 426 AlaLeuGluTygTygLysProSerThrThrLeuAapGlyThrAenLysLeuCysGlnPhe 445
 DB 483 GTTTTGAATACATCAAACTTCACATACCTGTGATGGACCAACAAAGCATATGCCATT 502
 QY 446 SerProValGlnLeuSerGluGlnThrProGluThrSerProAapLysGluGluAla 465
 DB 543 TCCCTCTTCAAGAACTATCGAGCAAGATCCCAAAACAGTCTGTGTAAGAGAGAGCC 602
 QY 466 SerLleProLysLysLeuGlnThrAlaArgProSerAapSerGlnSerLysArgLeuHis 485
 DB 603 AGCATCCCAAGAGCTGCAAGCCGAGGCTTCAAGACAGCCAGACCAAGCAATTCAT 662
 QY 486 SerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeuSerProLeuHisArg 505
 DB 663 TCGGTCAAGAACCAAGCAAGTGGCAGCCGCCAGAGAGTCCCTTTATCTCCACTGCATCGA 722

QY 506 SerGlySerValGluAapAenTyrgIsthSerPheLeu-PheGlyLeuSerThrSergI 525
 DB 723 AGTGGAGGCTGGAGAGCAATTCACACCAAGCTTCTTTTGGCTTCCACAGCCA 782
 QY 525 ngInHisLeuThryLysSerAlaGlyLeu-GlyLeuLysGlyTyrPHisSerAapLysLeu 545
 DB 783 GCAGCACTTCAGAAATCTGCTGCTGGGGGCTTAAGGGCTGCACTCGATATCTTGG 842
 QY 545 la-ProGlnThrSerThrPro 551
 DB 843 CCCCCCAGACCTTACCCCT 863

RESULT 3

BQ945892 883 bp mRNA linear EST 21-AUG-2002
 AGENCOURT 8926314 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6463045
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ945892
 VERSION BQ945892.1 GI:22361370
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Hierarchy: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 883)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>

Plate: LCM2653 row: 0 column: 14
 High quality sequence stop: 672.
 Location/Qualifiers

FEATURES

source

1..883
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6463045"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 9,34e-112 Length: 883
 Score: 1397.50 Matches: 286
 Percent Similarity: 97.29% Conservative: 1
 Best Local Similarity: 96.95% Mismatches: 8
 Query Match: 40.89% Indels: 3
 Gaps: 0

US-10-029-345A-109 (1-665) x BQ945892 (1-883)

QY 266 SerLeuApGluAlaTyrArgPheValIyGluLysArgProThrIleSerProAspPhe 285
 DB 3 TCTTTAGATGAGCTTACAGATTGTGAAAGAAAAAGACT-ACATATATCTCCAA--CTT 59
 QY 286 AenPheUenGlyGlnLeuLeuAapTyrgLylLysIleLysAenGlnThrgLylAser 305

```

Db      60  CATTTCCTGGGCAACTCTGAGTATGAGAAAGATTAGAACAGACGATGAGCATCA 119
Qy      306  GTPYLYSSELYSLEULYSLEULHISLEULYLSERPROANSGLUPROVALPROALA 325
Db      120  GGGCCAAAGACAACTTAAGCTGCTCACCTGAGAGCAATTAACCTGCTGCT 179
Qy      326  VALSERGLUGLYGILYLSERGLUTHRPROLEUSERPROPCYALASPSEALA 345
Db      180  GTCTCAGAGGGTGAAGAAAGAGAGACCCCTTCACTCAGTCACCTGTCACACTGCT 239
Qy      346  THRSERGLUALAAGLYGILNAGPROVALHISPROALASERVALPROSERVALPROSER 365
Db      240  ACCCTCAGAGGAGAGAGAGCAAAAGCCCGTGCATCCGCGCAGCGTCCAGCGTCCAGC 299
Qy      366  VALGILNPROSERLEULEUGLUNASPSERPROLEUVALGILNALALEUSERYLEULHISLEU 385
Db      300  GTGCAGCGCTGCTGTTAGAGGAGACCCGCTGCTGTAAGAGCGCTCAGAGGGCTGCACCTG 359
Qy      386  SERALASPARGLEUGLUNASPSERANLYSLEULYSARGSERPHESERLEUASPILALYS 405
Db      360  TCCGCAAGACAGCTGAGAGACAGCAATAAGCTCAAGCGTTCCTCTCTCGATATCAAA 419
Qy      406  SERVALSERLYSERALASERMERALALASERLEULHISGLYPHESERSERSERGLUASP 425
Db      420  TCAAGTTTCATATTTCAGCCAGATGCGACATCTTACATGAGCTTCTTCATCAGAAAGAT 479
Qy      426  ALALEUGLUTRYTYLYSPROSERTHRLEUASPGLYTHRANLYSLEUCYSGILNPE 445
Db      480  GTTTTGAATACATAAACCCTTCCACTCTGATGGAGCAACAGATATCCAGATTCC 539
Qy      446  SERPROVALINGILNLEUSERGLULNTHRPROGLNTHSERPROASGLUGLUALA 465
Db      540  TCCCTGTTTCAGAACTATCGAGACAGATCCGGAACAGATCTCATTAAGAGAGAGCC 599
Qy      466  SERILEPROLYSLEULNTHRALAARGPROSERASPSEGLSERLYSARGLEULHIS 485
Db      600  AGCATCCCAAGAGAGCTGACAGCCGCGCTTTCAGACAGCAGCAGAGAGCATTTGAT 659
Qy      486  SERVALARGTHRSESRSESGLYTHRALAGLARGSERLEULEUSERPROLEULHISARG 505
Db      660  TCGGTCAAGAAACAGACAGAGTGGCAACCGCCAGAGATCCCTTTATCTTCACTCATCGA 719
Qy      506  SERGLYSERVALGILUNASPSANTYRHSITHRSERPHELEUPHEGLYLEUSERTHSERGLN 525
Db      720  AGTGGAGAGCTGAGAGCAATTACCAACACACTTCTTTCGGGCTTTCACACAGCCAG 779
Qy      526  GLNHISLEUTHRILYSERVALAGLYLEUGLYLEULYSGLYTRPHISERASPILALEULA 545
Db      780  CAGCACTCAGCAAGAGCTGCTGCGCTGGGCTTAA-GGCTGGCACTGGATATCTGGGCC 838
Qy      546  PROGLNTHSERTHRPROSERLEUTHSERSERTPTPYRPHALA 560
Db      839  CCCCAAGCTCTAACCCCTTCCCTGACAGCAGCTGGATATTTGGC 883

```

RESULT 4
 LOCUS BQ951695 974 bp mRNA linear EST 21-AUG-2002
 DEFINITION AGENCOURT 8947130 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6461668
 ACCESSION BQ951695
 VERSION BQ951695.1 GI:22367173
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES

source

Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNC2650 row: F column: 05
 High quality sequence stop: 617.
 Location/Qualifiers
 1..974
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6461668"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_101"
 /note="Organ: lung; Vector: pOTB7; Site 1: Scori; Site 2:
 into Rcor1/Xho1 sites using the following 5' adaptor:
 GGCACAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Mismatches:	Indels:	Gaps:
9	376-110	974	1375-50	291	1	3
Percent Similarity:	91.82%					
Best Local Similarity:	91.51%					
Query Match:	40.24%					

US-10-029-345A-109 (1-665) x BQ951695 (1-974)

```

Qy      266  SERLEUASPGLUALAATYRARGPHEVALYSGILNLYSARGPROTHRILESERPROANPHE 285
Db      3  TCTTTAGATGAAGCTTACAGATTGTGTAAGAAAGAAAGAAAGCTTATCTCAAACTTC 62
Qy      286  AANPHELEUGLYGILNLEULASAPYRGLULYLYSILYLSANGLNTHRGYALASER 305
Db      63  AATTTTCGGGCCAACTCCCTGCACTATGAGAGAAAGATTAAACACAGCTGAGACATCA 122
Qy      306  GLYPROLYSSELYSLEULYSLEULHISLEUGLULYLSERPROANSGLUPROVALPROALA 325
Db      123  GGGCCAAAGAGCAAACTCAAGCTGCTGACACTGAGAGCCAAAGAACTGCTGCTGCT 182
Qy      326  VALSERGLUGLYGILNLYSSESGLYUTHRPROLEUSERPROPCYALASPSEALA 345
Db      183  GTCTCAGAGGGTGAAGAAAGAGAGAGCCCTTCACTCAGCTCAGCTGCGCCAGCTGCT 242
Qy      346  THRSERGLUALAAGLYGILNAGPROVALHISPROALASERVALPROSERVALPROSER 365
Db      243  ACCCTCAGAGGAGAGAGAGCAAAAGCCCGTGCATCCGCGCAGCGTCCAGCGTCCAGC 302
Qy      366  VALGILNPROSERLEULEUGLUNASPSERPROLEUVALGILNALALEUSERYLEULHISLEU 385
Db      303  GTGCAGCGCTGCTGTTAGAGGAGACCCGCTGCTGTAAGAGCGCTCAGAGGGCTGCACCTG 362
Qy      386  SERALASPARGLEUGLUNASPSERANLYSLEULYSARGSERPHESERLEUASPILALYS 405
Db      363  TCCGCAAGACAGCTGAGAGACAGCAATAAGCTCAAGCGTTCCTCTCTCGATATCAAA 422
Qy      406  SERVALSERLYSERALASERMERALALASERLEULHISGLYPHESERSERSERGLUASP 425
Db      423  TCAAGTTTCATATTTCAGCCAGATGCGACATCTTACATGAGCTTCTTCATCAGAAAGAT 482
Qy      426  ALALEUGLUTRYTYLYSPROSERTHRLEUASPGLYTHRANLYSLEUCYSGILNPE 445
Db      483  GTTTTGAATACATAAACCCTTCCACTCTGATGGAGCAACAGATATCCAGATTCC 542

```

QY 446 SerProValGlnGluSerGlnGlnThrProGluThrSerProAspGlnGluAla 465
Db 543 TCCCTGTTAGAGAACTATGAGACAGACTCCGAAACCACTCTGATTAAGAGAGGCC 602
QY 466 SerIleProValysLeuGlnThrAlaArgProSerAspSerGlnSerIleValGluHis 485
Db 603 AGCATCCCAAGAAAGCTGAGACCGCCAGGCTTCACAGACGACGAAAGCATTCGCAT 662
QY 486 SerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuSerProLeuHisArg 505
Db 663 TCGGTGAGAACAGACAGACAGTGGACCGCCAGAGGTCCCTTTATCTTCACATGCATCGA 722
QY 506 SerGlySerValGluAspAsnTrpHisThrSerPhe--LeuPheGlyLeuSerThrSer 524
Db 723 AGTGGAGCCGTGAGAGCAATTACACACACCGCTCCCTTTCCGCTTCACACAGC 782
QY 525 GlnGlnHis-LeuThrIysSerAlaGlyLeuGlyLeuLys-GlyTrpHisSerAspIle- 543
Db 783 CAGCAGACCCCTCAACGAATCTGCTGGCTGGGCTTCAGGCTGGGCACTTCGAAATAT 842
QY 544 -LeuAlaProGlnThrSerThrPro-SerLeu-ThrSerSerTrpTrpPheAlaThrGln 562
Db 843 CTTTGGGCCCCAGACCTTACCTTCCTTCCTGAAACAGAGCTGGTAATTTTGGCCACG 902
QY 562 uSerSerHis--PheTrpSerAlaSerAla-IleTrpGlyGly 575
Db 903 AAGCCTCAGACCTTCTACTGCTCAGCCATTAAAGAGG 946

RESULT 5
BO933499 859 bp mRNA linear EST 21-AUG-2002
LOCUS DEFINITION
AGENCOURT_8732489 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6455595
5', mRNA sequence.
BO933499
BO933499.1 GI:22348882
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 859)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LICM2634 row: i column: 04
High quality sequence, stop: 727.
Location/Qualifiers
1. 859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone IMAGE:6455595"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_101"
/note="Organ: lung; Vector: pOT7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGACGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
SuperScript II RT (Life Technologies). Note: this is a
NIH_MGC library."

FEATURES

source

1. 859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone IMAGE:6455595"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_101"
/note="Organ: lung; Vector: pOT7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGACGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
SuperScript II RT (Life Technologies). Note: this is a
NIH_MGC library."

ORIGIN

Alignment Scores:
Pred. No.: 8,376-110 Length: 859
Score: 1375.00 Matches: 279
Percent Similarity: 97.89% Conservative: 0
Best Local Similarity: 97.89% Mismatches: 2
Query Match: 40.23% Indels: 4
DB: 13 Gaps: 0

US-10-029-345A-109 (1-665) x BO933499 (1-859)

QY 266 SerLeuAspGlnAlaTrpArgPheValIysGluLysArgProThrIleSerProAspPhe 285
Db 3 TCTTATGATGAGACTTACAGATTTTGTGAAGAAAAGACCTTCTATCTCAAACTTC 62
QY 286 AsnPheLeuGlyGlnLeuLeuAspTrpGlyLysIleLysAsnGlnThrGlyAlaSer 305
Db 63 AATTTCTGGGCCAACTCCCTGGACTATGAGAAAGATTAAAGAACAGACTGGAGCATCA 122
QY 306 GlyProLysSerIysLeuLysLeuLeuHisLeuGluLysProAsnGluProValProAla 325
Db 123 GGCCCAAGAGCAAACTCAAGCTGCTGCACCTGGAGACCAAAATGAACCTGCTGCT 182
QY 326 ValSerGluGlyGlnLysSerGluThrProLeuSerProProGlyAlaAspSerAla 345
Db 183 GTCTCAGAGGTGGACAGAAAGCGAGCGCCCTCAGTCCACCTGTGCCACTTGTCT 242
QY 346 ThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerValProSerValProSer 365
Db 243 ACCTCAGAGGAGCAGAGCAAGAGCCGTCGATCCGCCAGGCTGCCAGCGCCAGC 302
QY 366 ValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeuSerGlyLeuHisLeu 385
Db 303 GTCCAGCCGTGCTGTGTAAGAGCAGCCCGCTGTGTACAGCGCTCAGTGGCTGCACCTG 362
QY 386 SerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPheSerLeuAspIleLys 405
Db 363 TCCGACAGAGGCTGGAAAGACAAATTAAGCTCAACGCTTCTTCTCTGATATCAAA 422
QY 406 SerValSerTrpSerAlaSerMetAlaAlaSerLeuHisGlyPheSerSerSerGluAsp 425
Db 423 TCACTTCAATATTCAGCCAGCATGGCAGCATCTTCAATGAGCTTCCATCATCAAGAT 482
QY 426 AlaLeuGlnTrpTrpLysProSerThrThrLeuAspGlyThrAsnLysLeuGlnPhe 445
Db 483 GTTTGGAATACTAACAACTTCCACTACTGTGATGGAGCAACAAAGCTATGCCAGTTC 542
QY 446 SerProValGlnGluSerGlnGlnThrProGluThrSerProAspGlnGluAla 465
Db 543 TCCCTGTTAGAGAACTATGAGACAGACTCCGAAACCAATCTGATTAAGAGAGGCC 602
QY 466 SerIleProValysLeuGlnThrAlaArgProSerAspSerGlnSerIleValGluHis 485
Db 603 AGCATCCCAAGAAAGCTGAGACCGCCAGGCTTCACAGACGACGAAAGCATTCGCAT 662
QY 486 SerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuSerProLeuHisArg 505
Db 663 TCGGTGAGAACAGACAGACAGTGGACCGCCAGAGGTCCCTTTATCTTCACATGCATCGA 722
QY 506 SerGlySerValGluAspAsnTrpHisThrSerPheLeuPheGlyLeuSerThrSerGln 525
Db 723 AGTGGAGCCGTGAGAGCAATTACACACACAGCTTCTTTGGCTTTCACACAGCCAG 782
QY 526 GlnHisLeuThrIysSerAlaGly-LeuGly-LeuLysGlyTrp-HisSerAspIleLeu 544
Db 783 CAGCAGCTCAGAAAGCTCTGCTGGCCCTTAAAGGCTGGANNCACTCGGATATTC 842
QY 544 uAlaPro 546
Db 843 TGGGCC 849

RESULT 6

B0721265
 LOCUS B0721265 898 bp mRNA linear EST 16-JUL-2002
 DEFINITION AGENCOURT 8294508 Lupski sympathetic_trunk Homo sapiens cDNA clone
 IMAGE:6194455 5', mRNA sequence.
 ACCESSION B0721265
 VERSION B0721265.1 GI:21860162
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 898)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-roma1.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM1359 row: 1 column: 08
 High quality sequence stop: 669.
 Location/Qualifiers
 1..898

FEATURES source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6194455"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Lupski_ssympathetic_trunk"
 /note="Vector: PCMV-SPORT (Life Technologies); Site_1:
 NotI; Site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCACGCGTCG-3' and
 5'-GACTAGTCTAGATCGGAGCGGCCGCTT(5)-3'. Size selected >
 1 kb for average insert length 1.9 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine); available through Life
 Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 2,74e-109 Length: 898
 Score: 1369.50 Matches: 279
 Percent Similarity: 95.27% Conservative: 3
 Best Local Similarity: 94.26% Mismatches: 12
 Query Match: 40.07% Indels: 4
 DB: 13 Gaps: 1

US-10-029-345a-109 (1-665) x B0721265 (1-898)

QY 4 GJWcttiegYThGnlllevalThGluArgLeuValAlaLeuLeuGluSerGlyThr 23
 DB 14 GAGATGATTGGAACCTCAATTTGTTACTGAGAGGTGGTGGCTGCTGGAAGTGGAAAG 73
 QY 24 GJLlyValLeuLeuLleasPserArgProPheValGluThrAsnThSerHisIleLeu 43
 DB 74 GAAAGAGTGGCTAATTGATTAGCGGCCCTTTGTGTGAATACAAATACATCCACATTTTG 133
 QY 44 GJLlAlleasnlleasncyserylsluMetlYsArgArgLeuGlnlAspLysVal 63
 DB 134 GAAGCCATTAAATCACTGCTCAAGCTTATGAGCGAAGGTTCCAAACGACAAAGTG 193
 QY 64 LeuIleThrgLuleuIleGlnHisSerAlaLysHlslyValAspIleAspCysSerGln 83

DB 194 TTAATTACAGAGCTATCCAGCATTCAGCGAAACATAAGTTGACATTGATTGACAGTCAG 253
 QY 84 LysValValValTyAspGlnSerSerGlnAspValAlaSerLeuSerAspCysPhe 103
 DB 254 AAGGTGTGATTGATGATCAAGAGCTCCCAAGATGTGGCTCTCTCTTCAAGACTGTTT 313
 QY 104 LeuThValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeuLeuAlaGly 123
 DB 314 CTCACCTGATCTCTGGGTAACTGAGAGAGCTTCAACTCTGTTCACCTGCTTGAAGGT 373
 QY 124 GlyPheAlaGluPheSerArgCysPheProGlyLeuGlnGlnLysSerThrLeuVal 143
 DB 374 GGGTGTGAGATCTCTCGTTGTTTCCCTGGGCTCTGTAAGAGAAATCCACTAGTC 433
 QY 144 ProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThrArgIleLeu 163
 DB 434 CTTACTGCAATTTCTGAGCCCTTCTTACCTGTGGCAACATTGGGCAACCGGAATTTCTT 493
 QY 164 ProAsnLeuTyLeuGluGlyCysGlnArgAspValLeuAsnLysGluLeuIleGlnlAsn 183
 DB 494 CCCAATCTTATCTTGGCTGCGCAGCGAGATGTCCTTCACAGAGAGCTGATGACGAGAT 553
 QY 184 GlyIleGlyTyValLeuAsnAlaSerTyThrCysProLysProAspPheIleProGlu 203
 DB 554 GGAATTTGTTATGTATTAATGCCAGCATATACCTGCTCCAAAGCTGACTTATCCCGAG 613
 QY 204 SerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeuProTyLeu 223
 DB 614 TCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
 QY 224 AspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysValLeuValHis 243
 DB 673 GACAAATCAAGATTTGATTAATGAGAAACAAAGCTTCAATGATGATGTTCTAGTCAC 732
 QY 244 CysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyIleMetLysArgMet 263
 DB 733 TGTTTAGTGGGATCTCCGCTCGGC-ACCATCGCTATCGCTTACATCAAGAAAGATG 791
 QY 264 AspMetSer-LeuAspGluAlaTyArgPheValLysGluLysArgProThrIleSerP 283
 DB 792 GACATGCTCTTAAAGTAACTTACAGATTTGTGAGAAAGAAAGAAAGCTTATATCTCC 851
 QY 283 AsnPheAsn---PheLeuGlyGlnLeuLeuAspTyArgLysLys 297
 DB 852 AAAATCTCAATTTTCTTGGCCAACTCCTGGCATTTATGAAAAA 897

RESULT 7
 LOCUS B0482429 903 bp mRNA linear EST 21-MAR-2001
 DEFINITION 602526595F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4650430 5',
 mRNA sequence.
 ACCESSION B0482429
 VERSION B0482429.1 GI:13414708
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 903)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-roma1.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM1431 row: m column: 23
 High quality sequence stop: 806.

FEATURES
SOURCE

```

location/Qualifiers
1. 903
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4650430"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NH MGC_21"
/node="Organ:placenta; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
downstream 5' adaptor: GGCAAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

```

ORIGIN

Alignment Scores:			
Pred. No.:	1,88e-108	Length:	903
Score:	1360.00	Matches:	281
Percent Similarity:	94.72%	Conservative:	6
Best Local Similarity:	92.74%	Mismatches:	10
Query Match:	39.79%	Indels:	6
DB:	12	Gaps:	2
US-10-029-345A-109 (1-665) x BG482429 (1-903)			
QY	172	ArgAspValIleuAnlYsgIleuIleGIInIaenGIYIleGIYValIleuAnIa	191
DB	3	CGAGATGTCCTCAACAAGAGCTGAAGCAGACAGAAATGGGATGTATGTATAATGCC	62
QY	192	SetTYrThCySerProIyProAspPheIIeProGIuSerHisIaPheIleuArValProVal	211
DB	63	AGCAATACCTGCTCAAAAGCCTGACTTATACCCCGAGTCTCATTTCTGCGTGTG	122
QY	212	AsnAspSerPheCyGslulYsIIeIeuProTfPLeuAspIySerValAspPheIIeGIu	231
DB	123	AAATGACAGCTTTGTGTGAGAAATTTTGCCTGTGCTGGACAAATCGTATTCATTGAG	182
QY	232	LysAlaIyAlaSerAnGIYCyValIleuValHisCyIeIeuIaGIYIleSerArgSer	251
DB	183	AAAGCAAAAGCCCTCAATGATGTGTCTTAAGCACTGTTTAGCTGGAGATCCCGCTCC	242
QY	252	AlaThrIleAlaIleIaIaIYrIIeMeIySargMeIaPmeSerIeuaAspIuaIaTYr	271
DB	243	GCCACCATCGCTATCGCTCAATCAATGAAGAAGATGACATGTCTTAGATGAAGCTTAC	302
QY	272	ArgPheValIyGslulYsArgProThrIleSerProaenPheAsnPheIeugIYIleu	291
DB	303	AGATTGTGTGAABAAAAGAAAGAACTTAATCTCCAACTTCATATTTCTGGCCAACTC	362
QY	292	LeuAspTYrGslulYsIleYsIleYsAnGIInThrIyAlaSerGIYProIySerIeIeu	311
DB	363	CTGAGACTATGAGAAGAAGATTAAAGACAGACTGAGCATCGGGCCAAAGACAACTC	422
QY	312	LysIeIeuIeuhIleIeuGIuIyProaenGIuProValProIaIaIaIeIeugIYGIIn	331
DB	423	AACTGTGTGACCTGGAAGAACCAATGAACCTGTCCGTGCTCTCGAGGGTGAAG	482
QY	332	LysSerGIuThrProIeuserProProCySAlaAspSerAlaThSerGIuAlaIaGIY	351
DB	483	AAAAGCGAGCGCCCTCACTCACTCCCTGTGCGACTTGTCTACTCTCAAGGACGACGA	542
QY	352	GlnArgProValHisProIaIaSerValProSerValProSerValGlnProSerIeIeu	371
DB	543	CAAAAGGCCGTGATCCCGCAGACGTGGCCACAGTGCACAGGATGCACCGCTGCTTA	602
QY	372	GlnAspSerProIeValGlnAlaIeuserGIYIeuhIleIeuserAlaAspArgIeGIu	391
DB	603	GAGACAGCCCGTGTGTACAGGGCTCATGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGT	662

Qy	AspergillusleuAArgSerPheSerleuAspIleIysSer-ValSerTyrSerAl	411
Db	663 GACAGCAGTAAAGCTCAAGCGTTCCTCTCTGTGATATCAATCAAGTTCAATTCAGC	722
Qy	411 aSerMetAlaAlaSerleuHis---GlyPheSerSerSerGluAsp-AlaLeuGluTyrT	430
Db	723 CAGCATGGCAGCAATCCCTTACATTGGCTTCACTCTCATCAAGAAATGCTTTGGAACT	782
Qy	430 YrlYsPro-SerThr-ThrIleuAsp---GlyThrAsnLysLeuGlyGlnPheSerProVa	448
Db	783 ACAAACCTTTCACATTACTGTGAATTGGACCAAGCAAGATTATGCCAGTTCTCCCTGT	842
Qy	448 IGIingLleuSerGluGlnThrProGluThrSerProAspLysGluGluAlaSerIlePr	468
Db	843 TCAGGAACATACGAGAGAGACTCCCGAACAACAGTCTCATTAAGGAGAGAACCATATCCC	902
Qy	468 Q 468	
Db	903 A 903	

LOCUS	BC038231	2207 bp	mRNA	linear	HTC 01-OCT-2002
DEFINITION	Homo sapiens similar to dual specificity phosphatase 8 clone				

FEATURES

```

/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5547764"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NTH MGC_72"

```


REFERENCE 1 (bases 1 to 769)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLM at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES
 Source
 Seq primer: pYX-5.
 Location/Qualifiers

1..769
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6406486"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to Bontalio, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with Not I and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is TGAGAGAGCC. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemlin Chan, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 6,06e-96 Length: 769
 Score: 1216.00 Matches: 229
 Percent Similarity: 94.09% Conservative: 10
 Best Local Similarity: 90.16% Mismatches: 15
 Query Match: 35.58% Indels: 0
 DB: 13 Gaps: 0

US-10-029-345A-109 (1-665) x B0704078 (1-769)

QY 93 GlnaPValAlaSerLeuSerSerAspCysPheLeuThrValLeuLeuGlyValLeuGlu 112
 DB 6 CAAGATGTTGGTCTCTCTCTGACAGCTGCTTCTCACTGACTCTCGGGTAAACCTGAG 65
 QY 113 LysSerPheAenSerValHisLeuLeuAlaGlyValPheAlaGluPheSerAryGysPhe 132
 DB 66 AGGAGCTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 125
 QY 133 ProGlyLeuGysGluGlyValSerThrLeuValProThrCysHisSerGlnProCysLeu 152
 DB 126 CTTGGCTCTGTATAGGAAGATCCATCTGATGCCATCCGATCATCTGAGCTTGCCTTA 185
 QY 153 ProValAlaAenIleGlyProThrArgIleLeuProAenLeuYrLeuGlyCysGlnArg 172
 DB 186 CTTGTTCTGGAACATTTGGGCACTCGAATCTTCCCAATCTTAATTTGAGCTGCAAGCA 245
 QY 173 AspValLeuAenIleGlyLeuLeuIleGlnAenGlyIleGlyTyrValLeuAenAlaSer 192
 DB 246 GATGTCTCAACAGAGCTGATGCAACAGAAATGGATGGCTAATGTGTTAATCCAGC 305

QY 193 TyrThrCysProIysProAspPheIleProGluSerHisPheLeuArgValProValAsn 212
 DB 306 AATACCTGTCCAAAGACCTGACTTATACCTGAAATCTCACTTCCGACAGGCTGGAAT 365
 QY 213 AspSerPheCysGluValIleLeuProTyrPheAspIysSerValAspPheIleGlyLeu 232
 DB 366 GACACCTTTTGTGAAAGAAATCTTCCATGCTGTGGACAGAGCTGTGATTTCAATTGAGAA 425
 QY 233 AlaIysAlaSerAenGlyCysValLeuValHisCysLeuAlaGlyIleSerAgsSerAla 252
 DB 426 GCAAAAGCTTCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485
 QY 253 ThrIleAlaIleAlaTyrIleMetIysArgMetAspMetSerLeuAspGluAlaTyrArg 272
 DB 486 ACTATTCCTATTTGCTTCACTCATGAAAGAGATGACATGTCTCTATATAGCTTACAGA 545
 QY 273 PheValIysGluIysArgProThrIleSerProAspPheAsnPheLeuGlyGlnLeu 292
 DB 546 TTTGTGAAGAAAGAAAGACCTACTATATCTCCGAATTTAATTTATGGGCCAATCATG 605
 QY 293 AspTyrGluIysIysIleIysAenGlnThrGlyAlaSerGlyProIysSerIysLeu 312
 DB 606 GACTATGAGAAAGCATTAATATACAGACTGGAATCTGAGGCCANAGACANACTGAG 665
 QY 313 LeuLeuHisIleGlnGluIysProAenGluProValProAlaValSerGluGlyValIys 332
 DB 666 CTCTGCTGCTTACAGAAACCTGAGTGCCTGCTGAGCTTCAAGAGGGGATGAG 725
 QY 333 SerGluThrProLeuSerProProCysAlaAspSerAlaThr 346
 DB 726 AGTGACTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767

RESULT 11
 LOCUS B1821804
 DEFINITION B1821804
 mRNA sequence.
 ACCESSION B1821804
 VERSION B1821804.1 GI:15933354
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLM at:
<http://image.llnl.gov>
 Plate: L1440 row: f column: 21
 High quality sequence stop: 805.
 Location/Qualifiers

FEATURES

source

1..836
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5176724"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is

destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	1e-92	Length:	836
Score:	1180.00	Matches:	260
Percent Similarity:	93.97%	Conservative:	5
Best Local Similarity:	92.20%	Mismatches:	15
Query Match:	34.52%	Indels:	9
DB:	12	Gaps:	0

US-10-029-345A-109 (1-665) x B1821804 (1-836)

```

Qy 50 CysSerIysLeuWetLysArgLeuGlnGlnAspIysValIleuIleThrGluLeuIle 69
Db 2 TGCTCCAGCTTATGAGCGAGAGGTTCCACACAGACAAAGCTTATATACAGAGCTCATC 61
Qy 70 GlnHisSerAlaIysHisIysValAspIleAspCysSerGlnIysValIValIYrAsp 89
Db 62 CAGCATTCAGCGAAACATTAAGTTGACATTGA-TGCAGTCAGAAAGCTTGATTACGAT 120
Qy 90 GlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrValIleuLeuGly 109
Db 121 CAAAGCTCCCAAGATGT-GCCTCTCTCTTCACAGCTGTTTTCACCTGACCTTCGTGGGT 179
Qy 110 LysLeuGlnIysSerPheAsnSerValHisIleuLeuAlaGlyIlePheAlaGluPheSer 129
Db 180 AAACGTGAGAAAGCTTCAACTCTGTTCACTGCTTCAGGTGGGTTTGCAGATTCTCT 239
Qy 130 ArgCysPheProGlyLeuGlyGlyIysSerThrLeuValProThrCysIleSerGln 149
Db 240 CGTAGTTTCCCTGGCCTCTGTGAAGAAATCCACTAGCCCTACCTGCAATTTCTCAG 289
Qy 150 ProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuIYrLeuGly 169
Db 300 CTTGCTTACCTGTTGCAACATTTGGGCAACCCGAATTTCTCCCAATCTTTATCTTGGC 359
Qy 170 CysGlnArgAspValIleAsnIysGluLeuIleGlnIleAsnGlyIleGlyIYrValIleu 189
Db 360 TGCCAGGAGATGTCTCTCAACAGAGAGCTATGACGAAATGGATCGGTTATGTGTTA 419
Qy 190 AsnAlaSerIYrThrCysProIysProAspPheIleProGluSerHisPheLeuArgVal 209
Db 420 AATGCCAGCAATACCTGTCCAAAGCTGACTTATCCCGAGTCTCATTTCCGTGCTG 479
Qy 210 ProValAsnAspSerPheCysGluIysIleLeuProThrIleuAspIysSerValAspPhe 229
Db 480 CCGTGTGAATACAGCTTCTGTGAGAAATTTTGGCGTGGCAATCAGATGATTTTC 539
Qy 230 IIGluIysAlaIysAlaSerAsnGlyCysValIleuValHisCysLeuAlaGlyIleSer 249
Db 540 ATTGAGAAAGCAAAAGCTTCAATGATGTGT-CTAAGTGACTGTTTAGCTGGGATCTCC 598
Qy 250 ArgSerAlaThrIleAlaIleAlaIYrIleWetIysArgMetAspMetSerLeuAspGlu 269
Db 599 CGCTCCGC-ACCATCGCTATCGC-TACATCATGAAAGAGATGAGCATGTCTTTAGATGAA 656
Qy 270 AlaIYrArgPheValIysGlu-LysArgProThrIleSerProAsnPheAsnPheLeu-G 289
Db 657 GCTTACGATTTTGGACAGAACAGAACCTTACATATCTTCCCAAACTTCAATTTTCTGGG 716
Qy 289 IyGlnLeuLeuAspIYrGluIysIysIleIysAsnGlnIthrGlyAlaSerGlyProIys 309
Db 717 GCCAGCTCTCTGACTATGAGAGAGATTTAAGAACGAGCTGGAGCATCGAGGCGC-AGA 775
Qy 309 exIysLeuIysLeuHisIleuGlnIysProAsnGluProValProAlaValSerGlu 329
Db 776 GCTTAATCAAGCTCTGCACTGTGG-AGAAGGCATATGAACTGCTGCTGTCTCAAGG 834

```

Qy 329 IY 329
Db 835 GG 836

RESULT 12

CF727177 715 bp mRNA linear EST 09-OCT-2003
LOCUS UI-M-HB0-ck-j-09-0-UT-r1 NIH_BMAP_HB0 Mus musculus cDNA clone
DEFINITION IMAGE:30548096 5', mRNA sequence.

ACCESSION CF727177 GI:37601345
VERSION CF727177.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 715)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabos-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouse1.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
Seq primer: pYX-5.
Location/Qualifiers
1..715
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30548096"
/issue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP HB0"
/note="Organ: Eye; Vector: pYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TATGTGAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Alignment Scores:

Pred. No.:	4.91e-90	Length:	715
Score:	1148.00 <td>Matches:</td> <td>223</td>	Matches:	223
Percent Similarity:	96.64%	Conservative:	7
Best Local Similarity:	93.70%	Mismatches:	8
Query Match:	33.59%	Indels:	2
DB:	14	Gaps:	0

US-10-029-345A-109 (1-665) x CF727177 (1-715)

```

Qy 87 ValIYrAspGlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrVal 106
Db 2 GTTATGATCAAAAGTTCCCAAGATGTGTTCTCTCTGCAACAGCTTTCACCTGTA 61
Qy 107 LeuLeuGlyIysSerPheAsnSerValHisIleuLeuAlaGlyIlePheAla 126

```

Db 62 CTTCTGGGTAACAGAGAGAGCTTCAACTCTGCTCCAGCTTGCGAGGCTTTCCT 121
|||
Qy 127 GluPheSerArgCysPheProGlyLeuGlySerGlyLeuValProThrCys 146
|||
Db 122 GAGTCTCTGCTGTTCTCCCTGCTGTAAGAGAAAGTCACTGATGCTTCACTGCG 181
|||
Qy 147 ILeSerGlnProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeu 166
|||
Db 182 ATATCTGAGCTTGTCTTCTGTTGCGAACAATGGGCGCAACTCGAATCTTCCCAATTC 241
|||
Qy 167 TyrlleuGlyCysGlnArgAspValIleuAsnIleuGlyIleGlnIleAsnGlyTleGly 186
|||
Db 242 TATCTGGCTCCGAGCGAGATGCTCTCAACAAGGACCTGATGCAACAAGATGGATGGC 301
|||
Qy 187 TyrlleuAsnAlaSerTyrThrCysProIleProAspPheIleProGlySerHisPhe 206
|||
Db 302 TATGTTAAATGCGAGAAATACCTGTCGAAGCCTGACTTCATACCTGAATCTCACTTC 361
|||
Qy 207 LeuArgValProValAsnAspSerPheCysGlyIleuValIleuProTyrLeuAspIleSer 226
|||
Db 362 CTGGAGAGCTCTGTAATGACAGCTTTTGTGAGAAATCCTACCATGGTTGGAAGAATCT 421
|||
Qy 227 ValAspPheIleGlyIleValIleAsnGlyCysValIleuValHisCysIleuAla 246
|||
Db 422 GTGCAATTCATGGAAGAAAGCAAAAGCCTCAATGGCTGTGCTTATCTCACTGCTTACCT 481
|||
Qy 247 GlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetIleAspMetAspMetSer 266
|||
Db 482 GGGATCTCGCTCGCGACATGCTATGCTATGCTTACATCAAGAGAGATGGAACATGCT 541
|||
Qy 267 LeuAspGlnAlaTyrArgPheValIleGlyIleAspProThrIleSerProAsnPheAsn 286
|||
Db 542 CTAGATGAGCTTACAGATTGTGAAAGAAAAAGACCTACTATATCTCCGAATTTTAT 601
|||
Qy 287 PheLeuGlyGlnLeuLeuAspTyrGlyIleValIleValIleValIleValIleValIle 306
|||
Db 602 TTTATGGGCAACCACTGAGCTATGAGAAAGCATTAATACAGACCTGGAATGTCAGGG 661
|||
Qy 307 ProIleSerIleLeuIleValIleLeuIleValIleValIleValIleValIleValIle 324
|||
Db 662 CCANAGAGCAA-CTGAGCTCTGCACTTACAGAAACCAAGT-GAGCCCGAGCTT 713
|||
RESULT 13 920 bp mRNA linear EST 20-OCT-2000
BE897795 601438457P1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3923247 5',
LOCUS mRNA sequence.
DEFINITION BE897795.1 GI:10363618
ACCESSION BE897795
VERSION BE897795
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 920)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTP/DRP
DNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LHAM9759 row: b column: 16
High quality sequence stop: 678.
Location/Qualifiers
1..920
source /organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3923247"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; site_1: NotI;
site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Alignment Scores:
Pred. No.: 7.36e-86 Length: 920
Score: 1102.50 Matches: 235
Percent Similarity: 89.30% Conservative: 7
Best Local Similarity: 86.72% Mismatches: 19
Query Match: 32.26% Indels: 11
DB: Gaps: 3

US-10-029-345A-109 (1-665) x BE897795 (1-920)

Qy 233 AlAluYAlaSer-AsnGlyCysValIleuValHisCysIleuAlaGlyTleSerArgSerAl 252
Db 2 GCAAAAGCCCTCCGGGTGAGATGTTCTTACGTGCACTGTTAAGCTGGAGATCTCCGCTCCG 61
Qy 252 aThrIleAlaIleAlaTyrIleMetIleAspMetAspMetSerIleuAspGlnAlaTyrArg 272
Db 62 CACCATCGCTATGCTTACATCATGAAGAGATGACATGCTTGTAGATGAACCTTACAG 121
Qy 272 gPheValIleGlyIleArgProThrIleSerProAsnPheAsnIleuGlyIleuLeu 292
Db 122 ATTGTGAAGAAAGAAAGAACTACTATATCTCAAACTTCAATTTCTGGGCAACTCCCT 181
Qy 292 uAspTyrGlyIleValIleValIleValIleValIleValIleValIleValIleValIle 312
Db 182 GGACTAGAAAGAAAGATTAAGAACAGATGAGATCAAGGCGCAAGAGCAAACTCA 241
Qy 312 sIleuLeuHisIleuGlyIleuProAsnGlnProValProAlaValSerGlyIleGlyIle 332
Db 242 GCTGCTGCACCTGGAAGAGCAATGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 301
Qy 332 sSerGlnThrProLeuSerProProCysAlaAspSerAlaThrSerGlnAlaIleGly 352
Db 302 AAGCGAGAGCGCCCTCAGTCCACCTGTCGCGCACTGCTACCTCAGAGGCAAGCAAGCA 361
Qy 352 uArgProValHisProAlaSerValProSerValProSerValGlnProSerIleuGly 372
Db 362 AAGGCCCGTGCATCCGCGC-----AGCGTGCACAGCGTGCACCGCTGCTTACG 412
Qy 372 uAspSerProLeuValGlnAlaIleuSerGlyLeuHisIleuSerAlaAspArgIleuGly 392
Db 413 GACACAGCCGCTGTAACAGCGCTCAGTGGCGTGCACCTGTCCGACAGACAGGTGAAGA 472
Qy 392 pSerAsnIleuLeuArgSerPheSerIleuAspIleValSerValSerTyrSerIle 412
Db 473 CACCAATTAAGCTCAACAGTTCCTCTCTGATATCAATCAAGTTTCATATTCACGCCG 532
Qy 412 tMetAlaIleSerLeuHisGlyPheSerSerSerGlnuAspAlaIleuGlyTyrTyrIle 432
Db 533 CATGGAGAGATCTTACATAGCTTCTCTCATCAGAAAGAGCTTGTGAATACCTACAAAC 592
Qy 432 oSerThrThrLeuAspGlyThrAsnIleuValIleuValIleuValIleuValIleuVal 452
Db 593 TTCCATCTACTCTGAGAGGACCAACAGCTATGCCGTTCTC-CTGTCCAGGAATATAC 651
Qy 452 rGlnGlnThrProGlyThrSerTyrAspIleGlyIleValIleSerIleProValIleuGly 472
Db 652 GGAGCAGACTCCCGAA---CAGTCTGATAGGAGCAACAGCATCCCAAGAGAGCTGCA 708
Qy 472 nThrAlaArgProSerAspSerGlnSerIleValIleuHisSerValArgThrSerSer 492
Db 709 GACCGAGGCGCTTACAGAAAGCCAGA-----GCAGCGATGCTTTCGTCAA 753

Qy 492 rGlyThrAlaGlnArgSerLeuSerPro 502
 Db 754 CCCACAGGTGGCAGCCAGGTCTTCTCC 784

RESULT 14
 BI816954/c
 LOCUS BI816954
 DEFINITION imagegc_10_2000/g1_410bdf41.x1 Soares_NPBMC Homo sapiens cDNA
 clone IMAGE:4140798 3', mRNA sequence.
 ACCESSION BI816954
 VERSION BI816954.1 GI:15911639
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 682)
 Kaleb, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and
 Prange, C.K.
 TITLE The I.M.A.G.E. Consortium quality control effort: clone
 resequencing for verification
 JOURNAL Unpublished (2001)
 COMMENT Other ESTs: BG058779
 Contact: Prange CK
 The I.M.A.G.E. Consortium
 Lawrence Livermore National Laboratory
 Livermore, CA, USA
 Email: help@image.llnl.gov
 This read has been verified (found to hit its original self in the
 correct orientation), as part of the I.M.A.G.E. Consortium quality
 control effort. High quality sequence is defined as having 100 or
 more base pairs with a phred quality value of 20 or greater, where
 a sliding window of 4 base pairs with a phred quality value of 15
 or greater marks the beginning and end of the sequence. For
 information on obtaining this clone, please contact
 info@image.llnl.gov, effort.
 Plate: LLM9388 row: k column: 7
 Seq primer: -21m13
 High quality sequence stop: 682.
 Location/Qualifiers
 1. 682
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4140798"
 /isue_type="lymphocyte"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="Soares NPBMC"
 /note="Organ: blood; Vector: pT73D-Pac; Site 1: NotI;
 Site 2: EcoRI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5'
 TGTACCAATCTGAAGTGGAGCGCGCGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified pT73 vector. Library
 is normalized; constructed in the laboratory of M. Bento
 Soares (University of Iowa)."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,12e-84 Length: 682
 Score: 1086.50 Matches: 214
 Percent Similarity: 96.89% Conservative: 4
 Best Local Similarity: 95.11% Mismatches: 6
 Query Match: 31.79% Indels: 2
 DB: 12 Gaps: 1
 US-10-029-345A-109 (1-665) x BI816954 (1-682)

Qy 49 AsnCysSerIysLeuWetLysArgArgLeuGlnGlnAspLysValLeuIleThrgLjLeu 68
 Db 673 AACGTNCCAGCTA--TAAAGCAAGGTCCAAAGAGCAAGAGTGT-ATTACAGAGCTC 618

FEATURES
 source

Qy 69 ILeGlnHisSerAlaLysHsLysValAspIleAspCysSerGlnLysValValTyr 88
 Db 617 ATCCAGAGATTACAGCAACCAATAGGTGACATTGATTCAGTACAGAAAGGTGAGTTTAC 558

Qy 89 AspGlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrValLeuLeu 108
 Db 557 GATCAAAAGTCCCAAGATGTGCTCTCTCTTCAAGAGTGTTCCTCACTGACTTCTG 498

Qy 109 GlyLysLeuGlnLysSerPheAsnSerValHisLeuLeuAlaGlyGlyPheAlaGluPhe 128
 Db 497 GGTAACTGAGAAAGCTTCAACTGTTCACTGCTGACAGTGGGTGTGAGTTTC 438

Qy 129 SerArgCysPheProGlyLeuCysGlnGlyLysSerThrLeuValProThrCysIleSer 148
 Db 437 TCTCGTGTGTTCTCTGCTCTGTGAAGAAATCACTTACTTCTTCACTTCACTTCT 378

Qy 149 GlnProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyrLeu 168
 Db 377 CAGCTTGCTTACCTGTGGCAACATTTGGGCCAACCGAATTCTTCCATCTTATCTT 318

Qy 169 GlyCysGlnArgAspValLeuAsnLysGluLeuIleGlnGlnAsnGlyIleGlyTyrVal 188
 Db 317 GGCTGCACAGAGATGCTCTCAACAAGAGCTGATCAGAGAAATGGATGTGTTATGTG 258

Qy 189 LeuAsnAlaSerThrThrCysProLysProAspPheIleProGluSerHisPheLeuArg 208
 Db 257 TTAAATGCCAGCAATACCTGTCCAAAGCCTGATTTATCCCAAGTCTATTCCTGGCT 198

Qy 209 ValProValAsnAspSerPheCysGlnLysIleLeuProThrLeuAspLysSerValAsp 228
 Db 197 GTCCCTGTGAATGACAGCTTTGTGTGAAATTTTCCGTGTGGAACAATCAGTAGAT 138

Qy 229 PheIleGlnLysAlaLysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIle 248
 Db 137 TTCAATTGAGAAAGCAAAAGCTTCATGATGATGTGTTCTAGTGCACGTTTAGCTGGGATC 78

Qy 249 SerArgSerAlaThrIleAlaIleAlaTyrIleMerLysArgSerAspMetSerLeuAsp 268
 Db 77 TCCGCTCCGCAACCATCGTATCGCTATCATCATGAGAGATGATGATGTATGAT 18

Qy 269 GluAlaTyrArgPhe 273
 Db 17 GAAGCTTACAGATT 3

RESULT 15
 CF532917 656 bp mRNA linear EST 12-SEP-2003
 LOCUS UI-M-GH0-cgw-n-02-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone
 IMAGE:30357529 5', mRNA sequence.
 DEFINITION CF532917 GI:34584885
 ACCESSION CF532917
 VERSION CF532917.1
 KEYWORDS Mus musculus (house mouse)
 SOURCE EST.
 ORGANISM Mus musculus

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 656)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouseefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: PYX-5.
 Location/Qualifiers

FEATURES

GenCore version 5.1.6
Copyright (c) 1993 - 2004, Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:19:48 ; Search time 57.0786 Seconds

(without alignments)
3289.124 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 3418

Sequence: 1 MAHEMIGTQVTERLVALL.....LKVSGSSSRSSSMELIEVS 665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubppaa/PCRUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*

13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3406	99.6	665	9 US-09-816-494-2	Sequence 2, Appl1
2	3406	99.6	665	9 US-09-964-277-2	Sequence 2, Appl1
3	3406	99.6	665	12 US-10-072-012-680	Sequence 680, Appl
4	3406	99.6	665	12 US-10-168-506-14	Sequence 14, Appl
5	3406	99.6	665	12 US-10-343-357-7	Sequence 7, Appl1
6	3406	99.6	665	15 US-10-377-072-26	Sequence 26, Appl1
7	3406	99.6	665	16 US-10-257-026-2	Sequence 2, Appl1
8	3406	99.6	665	16 US-10-648-593-240	Sequence 240, App
9	3406	99.6	665	16 US-10-648-593-247	Sequence 247, App
10	3406	99.6	665	12 US-10-072-012-679	Sequence 679, App
11	3406	99.6	665	12 US-10-072-012-703	Sequence 703, App
12	3406	99.6	665	12 US-10-425-114-54204	Sequence 54204, A
13	3399	99.4	665	12 US-10-072-012-681	Sequence 681, App
14	3399	99.4	665	15 US-10-094-749-2312	Sequence 2312, Ap
15	3379.5	98.9	662	12 US-10-072-012-258	Sequence 258, App

16	3358.5	98.3	680	12 US-10-072-012-256	Sequence 256, App
17	3079.5	90.1	660	12 US-10-072-012-682	Sequence 682, App
18	2930	85.7	672	12 US-10-296-115-1259	Sequence 1259, Ap
19	2721.5	79.6	677	12 US-10-072-012-683	Sequence 683, App
20	2500	73.1	517	9 US-09-964-277-21	Sequence 21, Appl
21	1326	38.8	625	12 US-10-072-012-699	Sequence 699, App
22	1302	38.1	663	12 US-10-072-012-700	Sequence 700, App
23	1297	37.9	253	15 US-10-108-260A-4872	Sequence 4872, Ap
24	1075.5	31.5	616	12 US-10-072-012-266	Sequence 266, App
25	917	26.8	501	12 US-10-072-012-702	Sequence 702, App
26	807	23.6	155	9 US-09-964-277-7	Sequence 7, Appl1
27	732.5	21.4	461	12 US-10-072-012-701	Sequence 701, Appl
28	654.5	19.1	169	14 US-10-346-356-15	Sequence 15, Appl
29	654.5	19.1	170	9 US-09-775-925-26	Sequence 26, Appl
30	654.5	19.1	170	9 US-09-847-519A-11	Sequence 11, Appl
31	654.5	19.1	170	12 US-10-655-073-16	Sequence 16, Appl
32	654.5	19.1	170	14 US-10-314-058-14	Sequence 14, Appl
33	654.5	19.1	170	14 US-10-405-008-16	Sequence 16, Appl
34	619	18.1	155	9 US-09-964-277-6	Sequence 6, Appl1
35	619	18.1	155	9 US-09-955-732-6	Sequence 6, Appl1
36	469	13.7	444	9 US-09-964-899-47	Sequence 47, Appl
37	469	13.7	482	12 US-10-058-270A-130	Sequence 130, App
38	469	13.7	482	14 US-10-346-356-2	Sequence 2, Appl1
39	469	13.7	482	16 US-10-648-593-164	Sequence 164, App
40	466.5	13.6	381	14 US-10-184-832-2	Sequence 2, Appl1
41	453	13.3	394	9 US-09-736-457-805	Sequence 805, App
42	453	13.3	394	9 US-09-902-941-805	Sequence 805, App
43	453	13.3	394	9 US-09-849-626-805	Sequence 805, App
44	453	13.3	394	12 US-10-283-017-805	Sequence 805, App
45	453	13.3	394	14 US-10-017-754-805	Sequence 805, App

ALIGNMENTS

RESULT 1	
US-09-816-494-2	
Sequence 2, Application US/09816494	
Patent No. US20020034807A1	
GENERAL INFORMATION:	
APPLICANT: Meyers, Rachel A.	
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY	
TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR	
FILE REFERENCE: 10448-030002	
CURRENT APPLICATION NUMBER: US/09/816,494	
PRIOR FILING DATE: 2001-03-23	
PRIOR APPLICATION NUMBER: US 60/191,858	
NUMBER OF SEQ ID NOS: 10	
SOFTWARE: FastSeq for Windows Version 4.0	
SEQ ID NO 2	
LENGTH: 665	
TYPE: PRT	
ORGANISM: Homo sapiens	
US-09-816-494-2	
Query Match	99.6%; Score 3406; DB 9; Length 665;
Best Local Similarity	Pred. No. 4.6e-250;
Matches 663; Conservative	1; Mismatches 1; Indels 0; Gaps 0;
QY	1 MAHEMIGTQVTERLVALLSGTEKVLIDSRFVENVNTSHLEAININCSKLMKRLQQ 60
DB	1 MAHEMIGTQVTERLVALLSGTEKVLIDSRFVENVNTSHLEAININCSKLMKRLQQ 60
QY	61 DKVLITELIHSKHKVDIDCSQKVVVYDQSSQDVASLSDDCFYVLVGLTKLSPNSVHL 120
DB	61 DKVLITELIHSKHKVDIDCSQKVVVYDQSSQDVASLSDDCFYVLVGLTKLSPNSVHL 120
QY	121 LAGGAFBFRSCFGLTEGKSTLVPTGISQPCLEVVANIGPRILPNLYLGGQRVNLKELI 180
DB	121 LAGGAFBFRSCFGLTEGKSTLVPTGISQPCLEVVANIGPRILPNLYLGGQRVNLKELI 180
QY	181 QQNGIGVLAASVTCRKPDRIPESHFLRVVNSFCCKILPMDKSVDFEAKASNGCV 240


```
Db 181 QONGIGYVLAASNTCPKEDFIPESHFLRPVNDSECFCEKILPWLDKSVDFIEKAKASNGCV 240
Qy 241 LVHCLAGISRSATTAIAIYIMKRMMSLDEAYRFYKEKRPPTISPNFNLGQLLDYEKKIKN 300
Db 241 LVHCLAGISRSATTAIAIYIMKRMMSLDEAYRFYKEKRPPTISPNFNLGQLLDYEKKIKN 300
Qy 301 QTGASGPKSKLKLHLEKPNPVPVAVSEGGOKSETPLSPCADSATSEAAQORPVHPASV 360
Db 301 QTGASGPKSKLKLHLEKPNPVPVAVSEGGOKSETPLSPCADSATSEAAQORPVHPASV 360
Qy 361 PSVPSVOPSLLEDBPVLQALSGHLMSADRLSDSNKLRSPSLDIKTSVYSASMAASLHGF 420
Db 361 PSVPSVOPSLLEDBPVLQALSGHLMSADRLSDSNKLRSPSLDIKTSVYSASMAASLHGF 420
Qy 421 SSSSDALEYKXPSTTLDTGINKLCOFSPVOELSEQTPETSPDKBEASIPKLTQARPSDSQ 480
Db 421 SSSSDALEYKXPSTTLDTGINKLCOFSPVOELSEQTPETSPDKBEASIPKLTQARPSDSQ 480
Qy 481 SKRLHSVRTSSSGTAQRSLSLPLHRSQSVEDNYHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
Db 481 SKRLHSVRTSSSGTAQRSLSLPLHRSQSVEDNYHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
Qy 541 SDILAPQSTPSTLSSWYFATESSHFYASAIYGSASYSAYSCSOLPTCGDQVYSVRRR 600
Db 541 SDILAPQSTPSTLSSWYFATESSHFYASAIYGSASYSAYSCSOLPTCGDQVYSVRRR 600
Qy 601 QKPSDRADSRMSWHEESPFEKQFKRRSCOMEFGESIMSENRSREELGKYGQSSFGSGME 660
Db 601 QKPSDRADSRMSWHEESPFEKQFKRRSCOMEFGESIMSENRSREELGKYGQSSFGSGME 660
Qy 661 IIEVS 665
Db 661 IIEVS 665

RESULT 2
US-09-964-277-2
/ Sequence 2, Application US/09964277
/ Patent No. US20020137170A1
/ GENERAL INFORMATION:
/ APPLICANT: Luche, Ralf M.
/ APPLICANT: Wei, Bo
/ TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
/ FILE REFERENCE: 200125.434
/ CURRENT APPLICATION NUMBER: US/09/964.277
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 665
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-964-277-2

Query Match 99.6%; Score 3406; DB 9; Length 665;
Best Local Similarity 99.7%; Pred. No. 4.6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAHMIQTQIVTERIVALLESGETEKVLLIDSRPFVEYNTSHLEAININCSKLMKRLQQ 60
Db 1 MAHMIQTQIVTERIVALLESGETEKVLLIDSRPFVEYNTSHLEAININCSKLMKRLQQ 60
Qy 61 DKVLTIELIHSAXHKVDIDCSQKVYVYDQSSQVAVLSDDCFVTLLGKLEKSPNSVHL 120
Db 61 DKVLTIELIHSAXHKVDIDCSQKVYVYDQSSQVAVLSDDCFVTLLGKLEKSPNSVHL 120
Qy 121 LAGGFAPSRCPGLGCGKSTLVPTCISOPCLPVANIGPRTIIPNYLGGORVLMKEEL 180
Db 121 LAGGFAPSRCPGLGCGKSTLVPTCISOPCLPVANIGPRTIIPNYLGGORVLMKEEL 180
Qy 181 QONIGIVLAASNTCPKEDFIPESHFLRPVNDSECFCEKILPWLDKSVDFIEKAKASNGCV 240
Db 181 QONIGIVLAASNTCPKEDFIPESHFLRPVNDSECFCEKILPWLDKSVDFIEKAKASNGCV 240
```

```
Db 181 QONGIGYVLAASNTCPKEDFIPESHFLRPVNDSECFCEKILPWLDKSVDFIEKAKASNGCV 240
Qy 241 LVHCLAGISRSATTAIAIYIMKRMMSLDEAYRFYKEKRPPTISPNFNLGQLLDYEKKIKN 300
Db 241 LVHCLAGISRSATTAIAIYIMKRMMSLDEAYRFYKEKRPPTISPNFNLGQLLDYEKKIKN 300
Qy 301 QTGASGPKSKLKLHLEKPNPVPVAVSEGGOKSETPLSPCADSATSEAAQORPVHPASV 360
Db 301 QTGASGPKSKLKLHLEKPNPVPVAVSEGGOKSETPLSPCADSATSEAAQORPVHPASV 360
Qy 361 PSVPSVOPSLLEDBPVLQALSGHLMSADRLSDSNKLRSPSLDIKTSVYSASMAASLHGF 420
Db 361 PSVPSVOPSLLEDBPVLQALSGHLMSADRLSDSNKLRSPSLDIKTSVYSASMAASLHGF 420
Qy 421 SSSSDALEYKXPSTTLDTGINKLCOFSPVOELSEQTPETSPDKBEASIPKLTQARPSDSQ 480
Db 421 SSSSDALEYKXPSTTLDTGINKLCOFSPVOELSEQTPETSPDKBEASIPKLTQARPSDSQ 480
Qy 481 SKRLHSVRTSSSGTAQRSLSLPLHRSQSVEDNYHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
Db 481 SKRLHSVRTSSSGTAQRSLSLPLHRSQSVEDNYHTSFLFGLSTSOOHLTKSAGLGLKGMH 540
Qy 541 SDILAPQSTPSTLSSWYFATESSHFYASAIYGSASYSAYSCSOLPTCGDQVYSVRRR 600
Db 541 SDILAPQSTPSTLSSWYFATESSHFYASAIYGSASYSAYSCSOLPTCGDQVYSVRRR 600
Qy 601 QKPSDRADSRMSWHEESPFEKQFKRRSCOMEFGESIMSENRSREELGKYGQSSFGSGME 660
Db 601 QKPSDRADSRMSWHEESPFEKQFKRRSCOMEFGESIMSENRSREELGKYGQSSFGSGME 660
Qy 661 IIEVS 665
Db 661 IIEVS 665

RESULT 3
US-10-072-012-680
/ Sequence 680, Application US/10072012
/ Publication No. US2004003493A1
/ GENERAL INFORMATION:
/ APPLICANT: Tchenerov, Velizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zerhusen, Bryan
/ APPLICANT: Patuturajan, Meera
/ APPLICANT: Shinkete, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangoli, Esha
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Miller, Charles B.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Colman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Karazyna
/ APPLICANT: Grosse, William M.
/ APPLICANT: Alsobrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072.012
/ PRIOR FILING DATE: 2002-01-31.
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
```


QY 601 QKPSDRADSRSSWHEESPPEKOFKRRSCOMERGESIMSENRSREELIGKVGSQSSPFGSME 660
DB 601 QKPSDRADSRSSWHEESPPEKOFKRRSCOMERGESIMSENRSREELIGKVGSQSSPFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 5

US-10-343-357-7
Sequence 7, Application US/10343357
Publication No. US20040058341A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom
APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
APPLICANT: YAO, Monique G.; BURFORD, Neil
APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
APPLICANT: GANDHI, Ameena R.; ARVITU, Chandra S.
APPLICANT: LEE, Ernestine A.; HAFALIA, April J. A.
APPLICANT: LU, Dzung Aina M.; TRIBOULEY, Catherine M.
APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
APPLICANT: YUE, Henry; WARREN, Bridget A.
APPLICANT: NGUYEN, Daniel B.; CHAMLA, Narinder K.
APPLICANT: KEARNEY, Liam
TITLE OF INVENTION: PROTEIN PHOSPHATASES
FILE REFERENCE: PI-0173 PCT
CURRENT APPLICATION NUMBER: US/10/343,357
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: PCT/US01/23716
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,679
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/223,272
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/224,309
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 60/226,728
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 60/229,254
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: US 60/231,366
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CD1
US-10-343-357-7

Query Match 99.6%; Score 3406; DB 12; Length 665;
Best Local Similarity 99.7%; Pred. No. 4,66-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHMEIGQVTEELVALLESGETEKVLLIDSRPFVEVNTSHIEANINOSKLMKRRLOQ 60
DB 1 MAHMEIGQVTEELVALLESGETEKVLLIDSRPFVEVNTSHIEANINOSKLMKRRLOQ 60
QY 61 DKVLITELIHSAGHKVDIDCSQVVVYDOSSDVASLSDDCELVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIHSAGHKVDIDCSQVVVYDOSSDVASLSDDCELVLLGKLEKSFNSVHL 120
QY 121 LAGGFARFSCFPGLECGKSTIVPTCISQPCLVANIGETRLILPNLYGCGORDVNTKELI 180
DB 121 LAGGFARFSCFPGLECGKSTIVPTCISQPCLVANIGETRLILPNLYGCGORDVNTKELI 180
QY 181 QONGIGVILNASYTCRKPPIFESHRLRPVVDVSFCEKILPWLKDSVDPIKAVASNGCV 240
DB 181 QONGIGVILNASYTCRKPPIFESHRLRPVVDVSFCEKILPWLKDSVDPIKAVASNGCV 240

QY 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAFRYKRPPTISPMNPLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAFRYKRPPTISPMNPLGQLLDYEKKIKN 300
QY 301 QTGASGPKSKLKLHLEKNEFPVAVSBGGQKSTPLSPPCADSATSEAGRPVHPASV 360
DB 301 QTGASGPKSKLKLHLEKNEFPVAVSBGGQKSTPLSPPCADSATSEAGRPVHPASV 360
QY 361 PSVPSVQPSLLEDBPLVQALSGHLISADRLSDNSKLRSPSLIKSVYSASMAASLHGP 420
DB 361 PSVPSVQPSLLEDBPLVQALSGHLISADRLSDNSKLRSPSLIKSVYSASMAASLHGP 420
QY 421 SSSSEALAEYRSTTLTGKTKLCOFSPVOELSTQTEPSPDKKEASIPKKLQTARPSDQ 480
DB 421 SSSSEALAEYRSTTLTGKTKLCOFSPVOELSTQTEPSPDKKEASIPKKLQTARPSDQ 480
QY 481 SKRLHSVRTSSSGTAQRLSLPLHRSQSVEDNYHTSFLKGLSTSOQHLTKSAGLKGWH 540
DB 481 SKRLHSVRTSSSGTAQRLSLPLHRSQSVEDNYHTSFLKGLSTSOQHLTKSAGLKGWH 540
QY 541 SDILAPQTSPTSLTSSWYPTATSSHFYSASAIYGSASAYSQSLPTCCDQVYSVRR 600
DB 541 SDILAPQTSPTSLTSSWYPTATSSHFYSASAIYGSASAYSQSLPTCCDQVYSVRR 600
QY 601 QKPSDRADSRSSWHEESPPEKOFKRRSCOMERGESIMSENRSREELIGKVGSQSSPFGSME 660
DB 601 QKPSDRADSRSSWHEESPPEKOFKRRSCOMERGESIMSENRSREELIGKVGSQSSPFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 6

US-10-377-072-26
Sequence 26, Application US/10377072
Publication No. US2004009501A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals Inc.
APPLICANT: Curtis, Rory A. J.
APPLICANT: Logan, Thomas Joseph
APPLICANT: Glucksmann, Maria A.
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, MiYoung
TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
FILE REFERENCE: MP103-0180NMIM
CURRENT APPLICATION NUMBER: US/10/377,072
CURRENT FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: US 09/895,860
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215,370
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/723,806
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 60/187,455
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/843,297
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199,801
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: US 09/861,801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205,508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/816,494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/815,419

PRIOR FILING DATE: 2001-03-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 665
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-377-072-26

Query Match 99.6%; Score 3406; DB 15; Length 665;
Best Local Similarity 99.7%; Pred. No. 4.6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTEVRLVALLSSTGTEKLLIDSRPFVYNTSHLEAININCKMKRLQ 60
DB 1 MAHEMIGTQVTEVRLVALLSSTGTEKLLIDSRPFVYNTSHLEAININCKMKRLQ 60
QY 61 DKVLITELIOMSAKHVIDIDCQKVVVYDQSSQDVASLSDDCFVLVLLGKLEKSPNSVHL 120
DB 61 DKVLITELIOMSAKHVIDIDCQKVVVYDQSSQDVASLSDDCFVLVLLGKLEKSPNSVHL 120
QY 121 LAGGFAFSCPCFGLCEGKSTLVPTCISQPCLPVANIIPTRILPNIYLGQQRVNLKELI 180
DB 121 LAGGFAFSCPCFGLCEGKSTLVPTCISQPCLPVANIIPTRILPNIYLGQQRVNLKELI 180
QY 181 QONGIGVNLASTYCPKPPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGVNLASTYCPKPPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLLDYEKKIKN 300
QY 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPSPCADSATSEAGORVHPASV 360
DB 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPSPCADSATSEAGORVHPASV 360
QY 361 PSVPSVQPSLLEDSPLVQALSGHLADRLSDNKLKRSFLDIKISVSASMAASLHGF 420
DB 361 PSVPSVQPSLLEDSPLVQALSGHLADRLSDNKLKRSFLDIKISVSASMAASLHGF 420
QY 421 SSSSEDALEYKPPSTTLDTGNKLCQFSPVOELSEQTPETSPDKKEASIPKLLQTPARPSDQ 480
DB 421 SSSSEDALEYKPPSTTLDTGNKLCQFSPVOELSEQTPETSPDKKEASIPKLLQTPARPSDQ 480
QY 481 SKRLHSVRRSSSGTARSLSPHRSQSVEDNYHTSFLFGLSTSOQHLTKSAGLGLKGMH 540
DB 481 SKRLHSVRRSSSGTARSLSPHRSQSVEDNYHTSFLFGLSTSOQHLTKSAGLGLKGMH 540
QY 541 SDLLAPQTSPTSLTSSWYFATESHFSASAIYGGASYSAYSCSQLPTCGDQVYSVRRR 600
DB 541 SDLLAPQTSPTSLTSSWYFATESHFSASAIYGGASYSAYSCSQLPTCGDQVYSVRRR 600
QY 601 QKPSDRADSRMSWHEESPPEKQFKRRSCQMEFGESIMSENRREBELKGVSSQSSFGSME 660
DB 601 QKPSDRADSRMSWHEESPPEKQFKRRSCQMEFGESIMSENRREBELKGVSSQSSFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 7
US-10-257-026-2
; Sequence 2, Application US/10257026
; Publication No. US2004006859A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10KDM5
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07

NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-257-026-2

Query Match 99.6%; Score 3406; DB 16; Length 665;
Best Local Similarity 99.7%; Pred. No. 4.6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTEVRLVALLSSTGTEKLLIDSRPFVYNTSHLEAININCKMKRLQ 60
DB 1 MAHEMIGTQVTEVRLVALLSSTGTEKLLIDSRPFVYNTSHLEAININCKMKRLQ 60
QY 61 DKVLITELIOMSAKHVIDIDCQKVVVYDQSSQDVASLSDDCFVLVLLGKLEKSPNSVHL 120
DB 61 DKVLITELIOMSAKHVIDIDCQKVVVYDQSSQDVASLSDDCFVLVLLGKLEKSPNSVHL 120
QY 121 LAGGFAFSCPCFGLCEGKSTLVPTCISQPCLPVANIIPTRILPNIYLGQQRVNLKELI 180
DB 121 LAGGFAFSCPCFGLCEGKSTLVPTCISQPCLPVANIIPTRILPNIYLGQQRVNLKELI 180
QY 181 QONGIGVNLASTYCPKPPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGVNLASTYCPKPPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLLDYEKKIKN 300
QY 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPSPCADSATSEAGORVHPASV 360
DB 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGGQKSETPSPCADSATSEAGORVHPASV 360
QY 361 PSVPSVQPSLLEDSPLVQALSGHLADRLSDNKLKRSFLDIKISVSASMAASLHGF 420
DB 361 PSVPSVQPSLLEDSPLVQALSGHLADRLSDNKLKRSFLDIKISVSASMAASLHGF 420
QY 421 SSSSEDALEYKPPSTTLDTGNKLCQFSPVOELSEQTPETSPDKKEASIPKLLQTPARPSDQ 480
DB 421 SSSSEDALEYKPPSTTLDTGNKLCQFSPVOELSEQTPETSPDKKEASIPKLLQTPARPSDQ 480
QY 481 SKRLHSVRRSSSGTARSLSPHRSQSVEDNYHTSFLFGLSTSOQHLTKSAGLGLKGMH 540
DB 481 SKRLHSVRRSSSGTARSLSPHRSQSVEDNYHTSFLFGLSTSOQHLTKSAGLGLKGMH 540
QY 541 SDLLAPQTSPTSLTSSWYFATESHFSASAIYGGASYSAYSCSQLPTCGDQVYSVRRR 600
DB 541 SDLLAPQTSPTSLTSSWYFATESHFSASAIYGGASYSAYSCSQLPTCGDQVYSVRRR 600
QY 601 QKPSDRADSRMSWHEESPPEKQFKRRSCQMEFGESIMSENRREBELKGVSSQSSFGSME 660
DB 601 QKPSDRADSRMSWHEESPPEKQFKRRSCQMEFGESIMSENRREBELKGVSSQSSFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 8
US-10-648-593-240
; Sequence 240, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26

D.Pet X

;; PRIOR APPLICATION NUMBER: 60/406,385
;; PRIOR FILING DATE: 2002-08-27
;; NUMBER OF SEQ ID NOS: 557
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 240
;; LENGTH: 665
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-648-593-240

Query Match 99.6%; Score 3406; DB 16; Length 665;
Best Local Similarity 99.7%; Pred. No. 4.6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVEYNTSHLEAININCSKLMRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVEYNTSHLEAININCSKLMRRLOQ 60
QY 61 DKVLITELIOMSAKHVDIDCSOKVYVYDSSQDVASLSDCFLLVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIOMSAKHVDIDCSOKVYVYDSSQDVASLSDCFLLVLLGKLEKSFNSVHL 120
QY 121 LAGGFARFCRCPGLCEGKSTLVPTCISQPCLPVANIPTRIILPMLYICGQDVLNKELI 180
DB 121 LAGGFARFCRCPGLCEGKSTLVPTCISQPCLPVANIPTRIILPMLYICGQDVLNKELI 180
QY 181 QONGIGVYVNASYTCRCPDPIPESHFLKVPVNDSCFKILPWLKSDVPLEKASNGCV 240
DB 181 QONGIGVYVNASYTCRCPDPIPESHFLKVPVNDSCFKILPWLKSDVPLEKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFYKERPTISPNFNLGQLDYEEKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFYKERPTISPNFNLGQLDYEEKIKN 300
QY 301 QTGASGPKSKLKLHLEKPNPVAVSEGGQKSTPLSPCADSATSEAAQRPVHPASV 360
DB 301 QTGASGPKSKLKLHLEKPNPVAVSEGGQKSTPLSPCADSATSEAAQRPVHPASV 360
QY 361 PSVSVQPSLLEDSPLVQALSGHLISADRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
DB 361 PSVSVQPSLLEDSPLVQALSGHLISADRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
QY 421 SSSDALLEYKPTSTLDTGNTKLCQFSPVQELSQTEPETSPPDKEASIPKQLQTAAPSDQ 480
DB 421 SSSDALLEYKPTSTLDTGNTKLCQFSPVQELSQTEPETSPPDKEASIPKQLQTAAPSDQ 480
QY 481 SKRLHSVRTSSSGTAQRSLSPILHRSQVYEDNYHTSFLGLSTSOQHLTKSAGLKGWH 540
DB 481 SKRLHSVRTSSSGTAQRSLSPILHRSQVYEDNYHTSFLGLSTSOQHLTKSAGLKGWH 540
QY 541 SDILAPQTSPTSLTSSMYFATESSHFYASAIYGGASAYSAYSCGOLPTCGDOYVYVRR 600
DB 541 SDILAPQTSPTSLTSSMYFATESSHFYASAIYGGASAYSAYSCGOLPTCGDOYVYVRR 600
QY 601 QKPSRADRRSRWHESPEKQFKRRSCOMFEGSINSERNRRELKGVGSSQSSFSGME 660
DB 601 QKPSRADRRSRWHESPEKQFKRRSCOMFEGSINSERNRRELKGVGSSQSSFSGME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 9

US-10-648-593-247
;; Sequence 247, Application US/10648593
;; Publication No. US20040106132A1
;; GENERAL INFORMATION:
;; APPLICANT: Bristol-Myers Squibb Company
;; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
;; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
;; FILE REFERENCE: D0273 NP

D. Rat

;; CURRENT APPLICATION NUMBER: US/10/648,593
;; CURRENT FILING DATE: 2003-08-26
;; PRIOR APPLICATION NUMBER: 60/406,385
;; PRIOR FILING DATE: 2002-08-27
;; NUMBER OF SEQ ID NOS: 557
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 247
;; LENGTH: 665
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-648-593-247

Query Match 99.6%; Score 3406; DB 16; Length 665;
Best Local Similarity 99.7%; Pred. No. 4.6e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVEYNTSHLEAININCSKLMRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVEYNTSHLEAININCSKLMRRLOQ 60
QY 61 DKVLITELIOMSAKHVDIDCSOKVYVYDSSQDVASLSDCFLLVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIOMSAKHVDIDCSOKVYVYDSSQDVASLSDCFLLVLLGKLEKSFNSVHL 120
QY 121 LAGGFARFCRCPGLCEGKSTLVPTCISQPCLPVANIPTRIILPMLYICGQDVLNKELI 180
DB 121 LAGGFARFCRCPGLCEGKSTLVPTCISQPCLPVANIPTRIILPMLYICGQDVLNKELI 180
QY 181 QONGIGVYVNASYTCRCPDPIPESHFLKVPVNDSCFKILPWLKSDVPLEKASNGCV 240
DB 181 QONGIGVYVNASYTCRCPDPIPESHFLKVPVNDSCFKILPWLKSDVPLEKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFYKERPTISPNFNLGQLDYEEKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFYKERPTISPNFNLGQLDYEEKIKN 300
QY 301 QTGASGPKSKLKLHLEKPNPVAVSEGGQKSTPLSPCADSATSEAAQRPVHPASV 360
DB 301 QTGASGPKSKLKLHLEKPNPVAVSEGGQKSTPLSPCADSATSEAAQRPVHPASV 360
QY 361 PSVSVQPSLLEDSPLVQALSGHLISADRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
DB 361 PSVSVQPSLLEDSPLVQALSGHLISADRLSDNKLKRSFSLDIKVSYSASMAASLHGF 420
QY 421 SSSDALLEYKPTSTLDTGNTKLCQFSPVQELSQTEPETSPPDKEASIPKQLQTAAPSDQ 480
DB 421 SSSDALLEYKPTSTLDTGNTKLCQFSPVQELSQTEPETSPPDKEASIPKQLQTAAPSDQ 480
QY 481 SKRLHSVRTSSSGTAQRSLSPILHRSQVYEDNYHTSFLGLSTSOQHLTKSAGLKGWH 540
DB 481 SKRLHSVRTSSSGTAQRSLSPILHRSQVYEDNYHTSFLGLSTSOQHLTKSAGLKGWH 540
QY 541 SDILAPQTSPTSLTSSMYFATESSHFYASAIYGGASAYSAYSCGOLPTCGDOYVYVRR 600
DB 541 SDILAPQTSPTSLTSSMYFATESSHFYASAIYGGASAYSAYSCGOLPTCGDOYVYVRR 600
QY 601 QKPSRADRRSRWHESPEKQFKRRSCOMFEGSINSERNRRELKGVGSSQSSFSGME 660
DB 601 QKPSRADRRSRWHESPEKQFKRRSCOMFEGSINSERNRRELKGVGSSQSSFSGME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 10

US-10-072-012-679
;; Sequence 679, Application US/10072012
;; Publication No. US20040033493A1
;; GENERAL INFORMATION:
;; APPLICANT: Tcherny, Velizar
;; APPLICANT: Spytek, Kimberly
;; APPLICANT: Zetunen, Bryan

```

APPLICANT: Patnrajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Beha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 679
LENGTH: 690
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-679

```

```

Query Match          99.64; Score 3406; DB 12; Length 690;
Best Local Similarity 99.74; Pred. No. 4.9e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MAHEMIGTQVTRVLVALLESGETEKVLLIDSRPFVEYNTSHILEAININSGKLMKRLQ 60
DB 26 MAHEMIGTQVTRVLVALLESGETEKVLLIDSRPFVEYNTSHILEAININSGKLMKRLQ 85
QY 61 DKVLITELLQHSKHKVDIDCSQKVVVVYDSSQDVASLSDDCTLVLLGLKLEKSPNSVHL 120
DB 86 DKVLITELLQHSKHKVDIDCSQKVVVVYDSSQDVASLSDDCTLVLLGLKLEKSPNSVHL 145
QY 121 LAGGFAFSCFPGLCGKSTLVPTCISOPCLPVANIGPRTIIPNTYLGCGQDVLANKEIL 180
DB 146 LAGGFAFSCFPGLCGKSTLVPTCISOPCLPVANIGPRTIIPNTYLGCGQDVLANKEIL 205
QY 181 QONGIGVLANASTCPRDFIPESHFLRVVNDSCFCEKILPMLDKSVDFTEKAKANGCV 240
DB 206 QONGIGVLANASTCPRDFIPESHFLRVVNDSCFCEKILPMLDKSVDFTEKAKANGCV 265

```

```

QY 241 LVHCLAGISRSATTAIAYIMKMDMSLDEAYRFEKKEPTISPNFNLQGLDYEEKIKXN 300
DB 266 LVHCLAGISRSATTAIAYIMKMDMSLDEAYRFEKKEPTISPNFNLQGLDYEEKIKXN 325
QY 301 QTGASGPKSKLKLHLEKNEPVPVAVSEGGQSETPLSPPCADSATSEAGQRPVHPASV 360
DB 326 QTGASGPKSKLKLHLEKNEPVPVAVSEGGQSETPLSPPCADSATSEAGQRPVHPASV 385
QY 361 PSVPSVQPSLDESPVLVQALSGHLADRLDESNKLRKSPSLDIKVSYSASMAASLHGF 420
DB 386 PSVPSVQPSLDESPVLVQALSGHLADRLDESNKLRKSPSLDIKVSYSASMAASLHGF 445
QY 421 SSEDALFEYKPESTTDGNNKLCQFSPVQBLSEQTPETPPDKEASIPKLTQPARSDSQ 480
DB 446 SSEDALFEYKPESTTDGNNKLCQFSPVQBLSEQTPETPPDKEASIPKLTQPARSDSQ 505
QY 481 SKRLHSVRTSSSGTAQRLSPILHRSQSVEDNHTSFLGLSTSQHLLTKSAGLKGWH 540
DB 506 SKRLHSVRTSSSGTAQRLSPILHRSQSVEDNHTSFLGLSTSQHLLTKSAGLKGWH 565
QY 541 SDILAQOTSTPSTSSWYFATESHSHFYASAIYGSASYSAYSCQLPTCGDQVSVRRR 600
DB 566 SDILAQOTSTPSTSSWYFATESHSHFYASAIYGSASYSAYSCQLPTCGDQVSVRRR 625
QY 601 OKPSRADSRRSWHESSPEKOPKRRSCOMEFCEISIMENSRHEELGKYGSSPFGSME 660
DB 626 OKPSRADSRRSWHESSPEKOPKRRSCOMEFCEISIMENSRHEELGKYGSSPFGSME 685
QY 661 IIEVS 665
DB 686 IIEVS 690

```

```

RESULT 11
US-10-072-012-703
Sequence 703, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zethusen, Bryan
APPLICANT: Patnrajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Beha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-01-31

```



```

; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 703
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-703

```

```

Query Match      99.6%; Score 3406; DB 12; Length 690;
Best Local Similarity 99.7%; Pred. No. 4.9e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSEGTKEKVLIDSRPVEYNTSHILEAININCKMKRRLOQ 60
DB 26 MAHEMIGTQIVTERLVALLSEGTKEKVLIDSRPVEYNTSHILEAININCKMKRRLOQ 85
QY 61 DKVLITTELIOHSAKHKVDIDCSQKVVYDOSQDVASLSSDCFLTVLGLKLEKSFNSVHL 120
DB 86 DKVLITTELIOHSAKHKVDIDCSQKVVYDOSQDVASLSSDCFLTVLGLKLEKSFNSVHL 145
QY 121 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGFRLIPNLNYGCGQDVANKELM 180
DB 146 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGFRLIPNLNYGCGQDVANKELM 205
QY 181 QONGIGVYLNASTCPRPDPFPESHFLRVVNDSEFCCKILPMWLDKSVDFIEKAKXNCGV 240
DB 206 QONGIGVYLNASTCPRPDPFPESHFLRVVNDSEFCCKILPMWLDKSVDFIEKAKXNCGV 265
QY 241 LVHCLAGISRSATTAIAIYIMKRDMSLDEAYRFVKEKPTISPNNFPGQLLDYEKKIKN 300
DB 266 LVHCLAGISRSATTAIAIYIMKRDMSLDEAYRFVKEKPTISPNNFPGQLLDYEKKIKN 325
QY 301 QTGASGPKSKLKLHLKEKNEPVPVAVSEGGQKSETPPLSPCADSATSEAAQRPVHPASV 360
DB 326 QTGASGPKSKLKLHLKEKNEPVPVAVSEGGQKSETPPLSPCADSATSEAAQRPVHPASV 385
QY 361 PSVPVQPSLLEDSPLVQALSGHLHSADRLSDSNLKRFSFLDIKSYVSASMAASLHGF 420
DB 386 PSVPVQPSLLEDSPLVQALSGHLHSADRLSDSNLKRFSFLDIKSYVSASMAASLHGF 445
QY 421 SSSSEDALEYKXSTLLDGTNKLCOFSPVOELSEQTPETSPDKEBASIPKQLQTPARPSDQ 480
DB 446 SSSSEDALEYKXSTLLDGTNKLCOFSPVOELSEQTPETSPDKEBASIPKQLQTPARPSDQ 505
QY 481 SKRLHSVTTSSSGTAQORSLSPLRSGSVEDNHTSTLFGELSTSQOHLTKSAGLGKGMH 540
DB 506 SKRLHSVTTSSSGTAQORSLSPLRSGSVEDNHTSTLFGELSTSQOHLTKSAGLGKGMH 565
QY 541 SDIILAPQTSPLSTSSWYFATESHFPYASAIYGGASYSAYSCQPLPTCGQOYVSARR 600
DB 566 SDIILAPQTSPLSTSSWYFATESHFPYASAIYGGASYSAYSCQPLPTCGQOYVSARR 625
QY 601 QKPSDRADSRRWHEBSPFEKQPKRRSCOMEFGESIMSENRSREBELGKYGQSSSFGSGME 660
DB 626 QKPSDRADSRRWHEBSPFEKQPKRRSCOMEFGESIMSENRSREBELGKYGQSSSFGSGME 685
QY 661 IIEVYS 665
DB 686 IIEVYS 690

```

```

RESULT 12
US-10-425-114-54204
; Sequence 54204, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO: 54204
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-10-425-114-54204

```

```

Query Match      99.6%; Score 3406; DB 12; Length 690;
Best Local Similarity 99.7%; Pred. No. 4.9e-250;
Matches 663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSEGTKEKVLIDSRPVEYNTSHILEAININCKMKRRLOQ 60
DB 26 MAHEMIGTQIVTERLVALLSEGTKEKVLIDSRPVEYNTSHILEAININCKMKRRLOQ 85
QY 61 DKVLITTELIOHSAKHKVDIDCSQKVVYDOSQDVASLSSDCFLTVLGLKLEKSFNSVHL 120
DB 86 DKVLITTELIOHSAKHKVDIDCSQKVVYDOSQDVASLSSDCFLTVLGLKLEKSFNSVHL 145
QY 121 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGFRLIPNLNYGCGQDVANKELM 180
DB 146 LAGFAEFSRCFPGLCGKSTLVPTCISQCLPVANIGFRLIPNLNYGCGQDVANKELM 205
QY 181 QONGIGVYLNASTCPRPDPFPESHFLRVVNDSEFCCKILPMWLDKSVDFIEKAKXNCGV 240
DB 206 QONGIGVYLNASTCPRPDPFPESHFLRVVNDSEFCCKILPMWLDKSVDFIEKAKXNCGV 265
QY 241 LVHCLAGISRSATTAIAIYIMKRDMSLDEAYRFVKEKPTISPNNFPGQLLDYEKKIKN 300
DB 266 LVHCLAGISRSATTAIAIYIMKRDMSLDEAYRFVKEKPTISPNNFPGQLLDYEKKIKN 325
QY 301 QTGASGPKSKLKLHLKEKNEPVPVAVSEGGQKSETPPLSPCADSATSEAAQRPVHPASV 360
DB 326 QTGASGPKSKLKLHLKEKNEPVPVAVSEGGQKSETPPLSPCADSATSEAAQRPVHPASV 385
QY 361 PSVPVQPSLLEDSPLVQALSGHLHSADRLSDSNLKRFSFLDIKSYVSASMAASLHGF 420
DB 386 PSVPVQPSLLEDSPLVQALSGHLHSADRLSDSNLKRFSFLDIKSYVSASMAASLHGF 445
QY 421 SSSSEDALEYKXSTLLDGTNKLCOFSPVOELSEQTPETSPDKEBASIPKQLQTPARPSDQ 480
DB 446 SSSSEDALEYKXSTLLDGTNKLCOFSPVOELSEQTPETSPDKEBASIPKQLQTPARPSDQ 505
QY 481 SKRLHSVTTSSSGTAQORSLSPLRSGSVEDNHTSTLFGELSTSQOHLTKSAGLGKGMH 540
DB 506 SKRLHSVTTSSSGTAQORSLSPLRSGSVEDNHTSTLFGELSTSQOHLTKSAGLGKGMH 565
QY 541 SDIILAPQTSPLSTSSWYFATESHFPYASAIYGGASYSAYSCQPLPTCGQOYVSARR 600
DB 566 SDIILAPQTSPLSTSSWYFATESHFPYASAIYGGASYSAYSCQPLPTCGQOYVSARR 625
QY 601 QKPSDRADSRRWHEBSPFEKQPKRRSCOMEFGESIMSENRSREBELGKYGQSSSFGSGME 660
DB 626 QKPSDRADSRRWHEBSPFEKQPKRRSCOMEFGESIMSENRSREBELGKYGQSSSFGSGME 685

```


Qy 661 IIEVS 665
Db 686 IIEVS 690

RESULT 13

US-10-072-012-681

; Sequence 681, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernov, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 681
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-072-012-681

Query Match 99.4%; Score 3399; DB 12; Length 665;
Best Local Similarity 99.5%; Pred. No. 1.6e-249;
Matches 662; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAHEMIGTQVTRLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRRLOQ 60
|||||

Db 1 MAHEMIGTQVTRLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRRLOQ 60
Qy 61 DKVLITELIOHSAKHKVDIDCSQKVVYVDOSSQDVASLSDDCLVTYLLKLEKSFNSYHL 120
Db 61 DKVLITELIOHSAKHKVDIDCSQKVVYVDOSSQDVASLSDDCLVTYLLKLEKSFNSYHL 120
Qy 121 LAGFAEFGRFCFGLCEGKSTLVPTCISQPCLPVANIIGFRLIPNLVLCQORVNLKELI 180
Db 121 LAGFAEFGRFCFGLCEGKSTLVPTCISQPCLPVANIIGFRLIPNLVLCQORVNLKELI 180
Qy 181 QONGIGYVINAASYCPKPDFIPESHFLRVVNDSECEKLLPWLDSKVDPIEKAKANGCV 240
Db 181 QONGIGYVINAASYCPKPDFIPESHFLRVVNDSECEKLLPWLDSKVDPIEKAKANGCV 240
Qy 241 LVHCLAGISRATIAAYIMKMDNSLDAVRFVKEKRTIPNPFLLQQLDYEEKIKN 300
Db 241 LVHCLAGISRATIAAYIMKMDNSLDAVRFVKEKRTIPNPFLLQQLDYEEKIKN 300
Qy 301 QTGASGPKSKLKLHLEKNEPVPVASEGQKSETPSPCADSATSEAAQORPVPAV 360
Db 301 QTGASGPKSKLKLHLEKNEPVPVASEGQKSETPSPCADSATSEAAQORPVPAV 360
Qy 361 PSVPSVQPSLLEDSPLVQALSGHLASADRLBDSNKLKRSFSLDKSVYSASMAASLHGF 420
Db 361 PSVPSVQPSLLEDSPLVQALSGHLASADRLBDSNKLKRSFSLDKSVYSASMAASLHGF 420
Qy 421 SSEDALLEYKSTLTDGNNKLCQSPVQELSEQTPETSPDKEASIPKKLOTARSDSQ 480
Db 421 SSEDALLEYKSTLTDGNNKLCQSPVQELSEQTPETSPDKEASIPKKLOTARSDSQ 480
Qy 481 SKRLHSVRTSSSGTAORSLSPLHRSQVEDNYHTSFLGLSTQOHLTKSAGLGKGMH 540
Db 481 SKRLHSVRTSSSGTAORSLSPLHRSQVEDNYHTSFLGLSTQOHLTKSAGLGKGMH 540
Qy 541 SDILAPQISTPSLTSWYFATSSSHFYASAIYGSASYSAYSCQLPTCGQVYSVRR 600
Db 541 SDILAPQISTPSLTSWYFATSSSHFYASAIYGSASYSAYSCQLPTCGQVYSVRR 600
Qy 601 QKSPDADSRGRSHESPEKQPKRRSCOMECESTMSNRRREELGXGSGSFGSGME 660
Db 601 QKSPDADSRGRSHESPEKQPKRRSCOMECESTMSNRRREELGXGSGSFGSGME 660
Qy 661 IIEVS 665
Db 661 IIEVS 665

RESULT 14

US-10-094-749-2312

; Sequence 2312, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 08435/0160
; CURRENT APPLICATION NUMBER: US/10/094,749

```

; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2312
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2312

```

Query Match

```

Best Local Similarity 99.4%; Score 3399; DB 15; Length 665;
Matches 662; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 MAHEMIGTQIVTERLVALLESSTGTEKVLIDSRPVEYNTSHILEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQIVTERLVALLESSTGTEKVLIDSRPVEYNTSHILEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIIOHSAKHKVDIDCSQKVVVYDSSQDVASLSDCFLTVLGLKLEKSFNSVHL 120
DB 61 DKVLITELIIOHSAKHKVDIDCSQKVVVYDSSQDVASLSDCFLTVLGLKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCGKSTLVPCTISOPCLPVANIGFRLIPNLVYAGORDVANKELI 180
DB 121 LAGGFAEFSRCFPGLCGKSTLVPCTISOPCLPVANIGFRLIPNLVYAGORDVANKELI 180
QY 181 OONGIGVYVNASYTCRCPDPIPESHFLRVVNDSCFCEKILPWLIDKSVDFIEKAKSNGCV 240
DB 181 OONGIGVYVNASYTCRCPDPIPESHFLRVVNDSCFCEKILPWLIDKSVDFIEKAKSNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRMMSLDEAVRPVKEKRPITISNPNFLGQLLDYEKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKRMMSLDEAVRPVKEKRPITISNPNFLGQLLDYEKIKN 300
QY 301 OTGASGPKSKLHLKPNPVPVAVSEGGQKSETPLSPCADSATSEAAQGPVHPASV 360
DB 301 OTGASGPKSKLHLKPNPVPVAVSEGGQKSETPLSPCADSATSEAAQGPVHPASV 360
QY 361 PSVPSVOPSLIEDSPVLQALSGHLNSADRLSDNSNKLKSFSLDIDVSYSASMAASLHGF 420
DB 361 PSVPSVOPSLIEDSPVLQALSGHLNSADRLSDNSNKLKSFSLDIDVSYSASMAASLHGF 420
QY 421 SSSSDALEYKRPSTLIDGTNKLCOESPVOELSEQTPETSPDKEEASIPKLLQTPARPSDQ 480
DB 421 SSSSDALEYKRPSTLIDGTNKLCOESPVOELSEQTPETSPDKEEASIPKLLQTPARPSDQ 480
QY 481 SKRLHSVTRTSSSGTAQRSLSPLRHSGSYEDNYHTSFLGLSTGQOHLTYSAGIGLKGWH 540
DB 481 SKRLHSVTRTSSSGTAQRSLSPLRHSGSYEDNYHTSFLGLSTGQOHLTYSAGIGLKGWH 540
QY 541 SDILAPQSTPLSTSWYATATSSHPYASAIYGSASYSAYSCSQPTCGDQVYSVRRR 600
DB 541 SDILAPQSTPLSTSWYATATSSHPYASAIYGSASYSAYSCSQPTCGDQVYSVRRR 600
QY 601 QRPSPRADSRSRHSESPFEKQFRRSCQMEFGHSINSEKRSBEIGKYGOSSEFSGSME 660
DB 601 QRPSPRADSRSRHSESPFEKQFRRSCQMEFGHSINSEKRSBEIGKYGOSSEFSGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

```

RESULT 15

```

; Sequence 258, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchearev, Velizar
; APPLICANT: Splet, Kimberly

```

```

; APPLICANT: Zerkusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Futrak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCES: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 258
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-258

```

Query Match

```

Best Local Similarity 98.9%; Score 3379.5; DB 12; Length 662;
Matches 660; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

```

```

QY 1 MAHEMIGTQIVTERLVALLESSTGTEKVLIDSRPVEYNTSHILEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQIVTERLVALLESSTGTEKVLIDSRPVEYNTSHILEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIIOHSAKHKVDIDCSQKVVVYDSSQDVASLSDCFLTVLGLKLEKSFNSVHL 120
DB 61 DKVLITELIIOHSAKHKVDIDCSQKVVVYDSSQDVASLSDCFLTVLGLKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCGKSTLVPCTISOPCLPVANIGFRLIPNLVYAGORDVANKELI 180
DB 121 LAGGFAEFSRCFPGLCGKSTLVPCTISOPCLPVANIGFRLIPNLVYAGORDVANKELI 180
QY 181 OONGIGVYVNASYTCRCPDPIPESHFLRVVNDSCFCEKILPWLIDKSVDFIEKAKSNGCV 240
DB 181 OONGIGVYVNASYTCRCPDPIPESHFLRVVNDSCFCEKILPWLIDKSVDFIEKAKSNGCV 240

```

```
Qy 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGOLLDEKKIXN 300
    |||||
Db 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGOLLDEKKIXN 300
    |||||
Qy 301 QTGASGPKSLKLIHLEKNEPVPAVSEGGQKSETPLSPCADSATSSEAQRPVHPASV 360
    |||||
Db 301 QTGASGPKSLKLIHLEKNEPVPAVSEGGQKSETPLSPCADSATSSEAQRPVHPA 358
    |||||
Qy 361 PSVPSVQPSLLEDSPLVQALSGHLGSADRLSDSNKLRSPFLDIKSVYSASMAASLHGF 420
    |||||
Db 359 -SVPSVQPSLLEDSPLVQALSGHLGSADRLSDSNKLRSPFLDIKSVYSASMAASLHGF 417
    |||||
Qy 421 SSEDALFYKPSSTTLDGTNKLCOFSPVOELSEOTPETSPDKEEASIPKLOTPARPSDSQ 480
    |||||
Db 418 SSEDALFYKPSSTTLDGTNKLCOFSPVOELSEOTPETSPDKEEASIPKLOTPARPSDSQ 477
    |||||
Qy 481 SKRLHSVRTSSSGTAQRSLSPHLRSGSVEDNYHTSFLFGLSTSOCHLTRKAGLKGWH 540
    |||||
Db 478 SKRLHSVRTSSSGTAQRSLSPHLRSGSVEDNYHTSFLFGLSTSOCHLTRKAGLKGWH 537
    |||||
Qy 541 SDILAPQTSPTSLTSSWYFATESSHFYASAGAIYGGASAYSACSQLPTCGDQVYSVRRR 600
    |||||
Db 538 SDILAPQTSPTSLTSSWYFATESSHFYASAGAIYGGASAYSACSQLPTCGDQVYSVRRR 597
    |||||
Qy 601 QKPSDRADSRSMHESPFEEKQFKRRSCOMEFGESIMSENRREBELKVGSSQSSFGSME 660
    |||||
Db 598 QKPSDRADSRSMHESPFEEKQFKRRSCOMEFGESIMSENRREBELKVGSSQSSFGSME 657
    |||||
Qy 661 IIEVS 665
    |||||
Db 658 IIEVS 662
```

Search completed: June 21, 2004, 13:24:57
Job time : 58.0786 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:20:38 ; Search time 22.6939 Seconds

(without alignments)
1512.797 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 3418
Sequence: 1 MAHEMIGTOIVTERLVALLE.....LCKVGSQSSFRSGSMELIIVS 665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/1aa/5A COMB.pep:*

2: /cgn2_6/ptodata/2/1aa/5B COMB.pep:*

3: /cgn2_6/ptodata/2/1aa/6A COMB.pep:*

4: /cgn2_6/ptodata/2/1aa/6B COMB.pep:*

5: /cgn2_6/ptodata/2/1aa/PCTUS COMB.pep:*

6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3406	99.6	665	US-09-816-494-2	Sequence 2, Appl1
2	654.5	19.1	170	US-09-544-716-14	Sequence 14, Appl1
3	654.5	19.1	170	US-09-557-921-15	Sequence 15, Appl1
4	654.5	19.1	170	US-09-564-357-17	Sequence 17, Appl1
5	654.5	19.1	170	US-09-619-380-16	Sequence 16, Appl1
6	469	13.7	482	US-09-557-921-2	Sequence 2, Appl1
7	453	13.3	394	US-09-702-705-805	Sequence 805, App
8	453	13.3	394	US-09-736-457-805	Sequence 805, App
9	453	13.3	394	US-09-614-1248-805	Sequence 805, App
10	453	13.3	394	US-09-671-325-805	Sequence 805, App
11	453	13.3	394	US-09-589-184-805	Sequence 805, App
12	445	13.0	394	US-08-530-290-23	Sequence 23, Appl1
13	445	13.0	394	US-09-702-705-827	Sequence 827, App
14	445	13.0	394	US-09-736-457-827	Sequence 827, App
15	445	13.0	394	US-09-614-1248-827	Sequence 827, App
16	445	13.0	394	US-09-671-325-827	Sequence 827, App
17	445	13.0	394	US-09-589-184-827	Sequence 827, App
18	439.5	12.9	395	US-08-990-379-5	Sequence 5, Appl1
19	433	12.7	367	US-08-990-379-6	Sequence 6, Appl1
20	425	12.4	314	US-09-371-6718-11	Sequence 11, Appl1
21	420	12.3	367	US-08-530-290-24	Sequence 24, Appl1
22	409.5	12.0	313	US-08-990-379-7	Sequence 7, Appl1
23	409	12.0	314	US-09-164-193-22	Sequence 22, Appl1
24	409	12.0	314	US-09-221-448A-22	Sequence 22, Appl1
25	393.5	11.5	393	US-08-990-379-4	Sequence 4, Appl1
26	377	11.0	302	US-09-702-705-806	Sequence 806, App
27	377	11.0	302	US-09-736-457-806	Sequence 806, App

28	377	11.0	302	US-09-614-1248-806	Sequence 806, App
29	377	11.0	302	US-09-671-325-806	Sequence 806, App
30	377	11.0	302	US-09-589-184-806	Sequence 806, App
31	366	10.7	397	US-08-990-379-8	Sequence 8, Appl1
32	346.5	10.1	168	US-09-544-716-13	Sequence 13, Appl1
33	346.5	10.1	168	US-09-557-921-13	Sequence 13, Appl1
34	346.5	10.1	168	US-09-564-357-16	Sequence 16, Appl1
35	346.5	10.1	168	US-09-619-380-15	Sequence 15, Appl1
36	338.5	9.9	170	US-09-544-716-12	Sequence 12, Appl1
37	338.5	9.9	170	US-09-557-921-12	Sequence 12, Appl1
38	338.5	9.9	170	US-09-564-357-15	Sequence 15, Appl1
39	338.5	9.9	170	US-09-619-380-14	Sequence 14, Appl1
40	319	9.3	169	US-09-544-716-16	Sequence 16, Appl1
41	319	9.3	169	US-09-557-921-17	Sequence 17, Appl1
42	319	9.3	169	US-09-564-357-19	Sequence 19, Appl1
43	319	9.3	169	US-09-619-380-18	Sequence 18, Appl1
44	312	9.1	168	US-09-544-716-15	Sequence 15, Appl1
45	312	9.1	168	US-09-557-921-16	Sequence 16, Appl1

ALIGNMENTS

RESULT 1
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-494-2

Query Match	Score	DB 4;	Length	665;
Best Local Similarity	99.6%	Pred. No. 4.8e-288;		
Matches	663;	Conservative	1;	Mismatches 1;
			Indels	0;
			Gaps	0;
QY	1	MAHEMIGTOIVTERLVALLESGTEKYLIDSRPFVEYNTSHIDEAININCKLMKRLOQ	60	
DB	1	MAHEMIGTOIVTERLVALLESGTEKYLIDSRPFVEYNTSHIDEAININCKLMKRLOQ	60	
QY	61	DKVLITELIQSHAKHYVDIDCSQKVVVYDQSSQDVASLSQDCELTVTLGTEKSFNSVHL	120	
DB	61	DKVLITELIQSHAKHYVDIDCSQKVVVYDQSSQDVASLSQDCELTVTLGTEKSFNSVHL	120	
QY	121	LACGPAEFCRCPGCLCEGKSTLVPTCISQPCLEPVANIGPRILPNIYLGGQRYVNLKELI	180	
DB	121	LACGPAEFCRCPGCLCEGKSTLVPTCISQPCLEPVANIGPRILPNIYLGGQRYVNLKELI	180	
QY	181	QONGIGVYVNASATYCKPDPFIPESHFLRVVNDSPFCCKILPMLDKSVDFTEKAKANGCV	240	
DB	181	QONGIGVYVNASATYCKPDPFIPESHFLRVVNDSPFCCKILPMLDKSVDFTEKAKANGCV	240	
QY	241	LVMCLAGISRSATIAIAYIMKMDMSLDEAYRVEKRPPTISPNFNLGQLDYKKIKN	300	
DB	241	LVMCLAGISRSATIAIAYIMKMDMSLDEAYRVEKRPPTISPNFNLGQLDYKKIKN	300	
QY	301	QTASGPKSKLKLHLEKREPEVPAYSEGQKSETPLSPCADSATSEAGORPVHPASV	360	
DB	301	QTASGPKSKLKLHLEKREPEVPAYSEGQKSETPLSPCADSATSEAGORPVHPASV	360	
QY	361	PSVPSVQPSLLEDSPLVQALSGHLISADRLDSDNKLKRSFSLDKSVSYSSAASALHGF	420	

```
Db 361 PSVPSVQPSLIEDSPVQHLISADRLSDSNLKRSPFLDIKSYISASMAASLHCP 420
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 421 SSSSEDLLEYKSGSTTLIDGNTKLCQPSPVQELSEQPEPSPKERASTPKQLQIARPSDQ 480
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 421 SSSSEDLLEYKSGSTTLIDGNTKLCQPSPVQELSEQPEPSPKERASTPKQLQIARPSDQ 480
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 481 SKRLHSVTRSSSGTQORSLSPILRSGSVENHTSTLFGILSTSQOHLTKSAGLKGWH 540
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 481 SKRLHSVTRSSSGTQORSLSPILRSGSVENHTSTLFGILSTSQOHLTKSAGLKGWH 540
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 541 SDIILAPQSTSTSLSSWTFATESHPSASAIYGSGSASYSAGSCQLPTCGDQVYSVRR 600
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 541 SDIILAPQSTSTSLSSWTFATESHPSASAIYGSGSASYSAGSCQLPTCGDQVYSVRR 600
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 601 QKPSDRADSRSRWHEESPFEKOPFRKRSQCMERESITMSNRSREBLGKVSQSFSGSGME 660
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 601 QKPSDRADSRSRWHEESPFEKOPFRKRSQCMERESITMSNRSREBLGKVSQSFSGSGME 660
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 661 IIEVS 665
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 661 IIEVS 665
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15
```

```
RESULT 2
US-09-544-716-14
; Sequence 14, Application US/09544716
; Patent No. 6492157
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; TITLE OF INVENTION: DSP-9 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.415
; CURRENT APPLICATION NUMBER: US/09/544.716
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-544-716-14
```

```
Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 134 GLCEGK-STLVPTCISQPCLPVANIIGPTRILPNLYLGCORVYLNKELIQONGIGYVLNAS 192
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-544-716-14

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 193 YTCPPDPFIPESHFLRPVNDSPCEKILPWLKSVDFTEKAKASNGCVLVHCLAGISRSA 252
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-544-716-14

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 253 TTAIAIYMKRMDMSLDEAYRFYKERRPTISPNNFLGQLDYEEKIK 299
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-544-716-14

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 121 TTAIAIYMKTMGSSDDAYRFVKDRRPSISPNFNLGQLLEYERTLK 167
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-544-716-14
```

```
RESULT 3
US-09-557-921-15
; Sequence 15, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.416
; CURRENT APPLICATION NUMBER: US/09/557.921
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15
```

```
Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 134 GLCEGK-STLVPTCISQPCLPVANIIGPTRILPNLYLGCORVYLNKELIQONGIGYVLNAS 192
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 193 YTCPPDPFIPESHFLRPVNDSPCEKILPWLKSVDFTEKAKASNGCVLVHCLAGISRSA 252
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 253 TTAIAIYMKRMDMSLDEAYRFYKERRPTISPNNFLGQLDYEEKIK 299
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 121 TTAIAIYMKTMGSSDDAYRFVKDRRPSISPNFNLGQLLEYERTLK 167
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-557-921-15
```

```
RESULT 4
US-09-564-357-17
; Sequence 17, Application US/09564357
; Patent No. 6645753
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; TITLE OF INVENTION: DSP-5 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.413
; CURRENT APPLICATION NUMBER: US/09/564.357
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-564-357-17
```

```
Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 134 GLCEGK-STLVPTCISQPCLPVANIIGPTRILPNLYLGCORVYLNKELIQONGIGYVLNAS 192
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-564-357-17

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 193 YTCPPDPFIPESHFLRPVNDSPCEKILPWLKSVDFTEKAKASNGCVLVHCLAGISRSA 252
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-564-357-17

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 253 TTAIAIYMKRMDMSLDEAYRFYKERRPTISPNNFLGQLDYEEKIK 299
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-564-357-17

Query Match
Best Local Similarity 19.1%; Score 654.5; DB 4; Length 170;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

Db 121 TTAIAIYMKTMGSSDDAYRFVKDRRPSISPNFNLGQLLEYERTLK 167
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-564-357-17
```

```
RESULT 5
US-09-619-380-16
; Sequence 16, Application US/09619380
; Patent No. 6649391
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; TITLE OF INVENTION: DSP-11 DUAL SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.418
; CURRENT APPLICATION NUMBER: US/09/619.380
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
```

LENGTH: 170
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-619-380-16

Query Match 19.1%; Score 654.5; DB 4; Length 170;
 Best Local Similarity 72.5%; Pred. No. 3.6e-49;
 Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;

QY 134 GLCEGK-STLVPTCISQPCLPVANIGPTRLIPNLYLGCQDVLNKLQNGIGYVLMAS 192
 DB 1 GLCEGKRALLPMSLSQPCLPVPSVGLTRILPHLYLGSQDVLNKLQNGIGYVLMAS 60
 QY 193 YTCPKPPIESHPRLRPVNDSPCEKLLPWLKDSVDTEKAKSNGCVLVHCLAGISRSA 252
 DB 61 NSCPKPPICESRMRVPINDNYCEKLLPWLKDSIEFDKAKLSQCVIVHCLAGISRSA 120
 QY 253 TIAIAYIMKMDSLDEAYRFVKEKPTISPNFNLQQLDYEEKIK 299
 DB 121 TIAIAYIMKMGSSDADYRFVADRPISPNFNLQQLDYERTLK 167

RESULT 6
 US-09-557-921-2
 Sequence 2, Application US/09557921
 Patent No. 6551810
 GENERAL INFORMATION:
 APPLICANT: Luche, Ralf M.
 APPLICANT: Wei, Bo
 TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
 FILE REFERENCE: 200125.416
 CURRENT APPLICATION NUMBER: US/09/557.921
 CURRENT FILING DATE: 2000-04-20
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2
 LENGTH: 482
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-557-921-2

Query Match 13.7%; Score 469; DB 4; Length 482;
 Best Local Similarity 35.9%; Pred. No. 3.1e-33;
 Matches 107; Conservative 65; Mismatches 96; Indels 30; Gaps 7;

QY 27 LLDISRPVEYNTSHILEAININCS-KLMKRLQODKVLITELIO-HSAKHVLDICSOX 84
 DB 173 VIIDCRPFMEYNHSHIGAVHINCADKISRRLQOGKITVLDLISCEGKDSFRIPIFSKE 232
 QY 85 VVVYDQSSQDVASLSDCFLTVLIGKLEKSFNSVHLLAGCPAFPSRCPPGLCEBGK---- 139
 DB 233 IIVYDENTNEPSRMPQPLHIVLESIKRGEKELVLKGLSFFKQNHENICDNLQLOE 292
 QY 140 -----STLVPTCISQPCLP-VANIGPTRLIPNLYLGCQDVLNKLQNGIGY 187
 DB 293 CREVGASASASLSLPPPI--PTTPDLENABLTFLPFLPLGNEQDQDLDYTMQRLNIGY 350
 QY 188 VLNAS-----YTCPKPPIESHPRLRPVNDSPCEKLLPWLKDSVDTEKAKSNGCVLV 242
 DB 351 VIVVTHLPPLHYEKGIF---NYKRLPATDSNMQNROYFEERFEETBEHOGCKLLI 406
 QY 243 HCLAGISRSATIAIAYIMKMDSLDEAYRFVKEKPTISPNFNLQQLDYEEKIK 300
 DB 407 HCQAGVSRSAIYIAYIMKMTMTDAVYFVKCKRPIISPNLNFQQLLEFEDLNN 464

RESULT 7
 US-09-702-705-805
 Sequence 805, Application US/09702705
 Patent No. 6504010
 GENERAL INFORMATION:
 APPLICANT: Wang, Tonglong
 APPLICANT: Bangur, Chaitanya S.

APPLICANT: Lodes, Michael A.
 APPLICANT: Fanger, Gary
 APPLICANT: Vedvick, Tom
 APPLICANT: Carter, Darlick
 APPLICANT: Retter, Marc
 APPLICANT: Mannon, Jane
 APPLICANT: Fan, Liqun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 FILE REFERENCE: 210121.478C14
 CURRENT APPLICATION NUMBER: US/09/702.705
 CURRENT FILING DATE: 2000-10-30
 NUMBER OF SEQ ID NOS: 1833
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 805
 LENGTH: 394
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-702-705-805

Query Match 13.3%; Score 453; DB 4; Length 394;
 Best Local Similarity 31.4%; Pred. No. 5.3e-31;
 Matches 120; Conservative 73; Mismatches 143; Indels 46; Gaps 13;

QY 19 LBSGTEKVLIDSRPVEYNTSHILEAININCSKLMKRLQODKVLITELI--OHSAKHK 76
 DB 39 LPSG-GKCLHLDPRPLASAGYILGSVNRCTIYRR-AKGSVSLQILPABEVRAR 96
 QY 77 VDIIDCSQKVVYDQSSQDVASLSDCFLTVLIGKLEKSF--NSVHLLAGCPAFPSRCFPG 134
 DB 97 LRSGLVSAYIVYERBPRAESLREDSVLSLVQALRNARERTICLLKGVERFSSYPE 156
 QY 135 LCEGKSTL-----VPTCISQPC---CLPVA-----NIGTRILPNIYLGCQDVLNKL 179
 DB 157 FCSKTRALAIIPPVPSATEPPLDCCSCGTPRHQEGVEILPFLYLGSAYHARM 216
 QY 180 IQONGIGYVLMASYSYTCPKPPIESHPRLRPVNDSPCEKLLPWLKDSVDTEKAKSNGC 239
 DB 217 LDALGITLALNVSDDCN-HFEGHYQKCLPVEDNHRADISSPMETAIETIDAVKCRGR 275
 QY 240 VLHCLAGISRSATIAIAYIMKMDSLDEAYRFVKEKPTISPNFNLQQLDYEEKIK 299
 DB 276 VLHCLAGISRSATICAYIMKMRVLEAFEFVQRRATISPNFNLQQLDYEEKIK 335
 QY 300 NQTAGSPYSKLLHLKPNNEPVAVBSGQKSETPSPCADSATSEAGRPVHPAS 359
 DB 336 ATSCAELAAS-----PSGFL-----GERGKTPATP-----TSQVFSFPV-SVG 373
 QY 360 VPSVPSVQPSLSDPLVQALS 381
 DB 374 VHSAPSSLPYL--HSPITTSPS 393

RESULT 8
 US-09-736-457-805
 Sequence 805, Application US/09736457
 Patent No. 6509448
 GENERAL INFORMATION:
 APPLICANT: Wang, Tonglong
 APPLICANT: Bangur, Chaitanya S.
 APPLICANT: Lodes, Michael A.
 APPLICANT: Fanger, Gary
 APPLICANT: Vedvick, Tom
 APPLICANT: Carter, Darlick
 APPLICANT: Retter, Marc
 APPLICANT: Mannon, Jane
 APPLICANT: Fan, Liqun
 APPLICANT: Wang, Aijun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 FILE REFERENCE: 210121.478C15
 CURRENT APPLICATION NUMBER: US/09/736.457
 CURRENT FILING DATE: 2000-12-13


```

; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-805
```

```

Query Match          13.3%; Score 453; DB 4; Length 394;
Best Local Similarity 31.4%; Pred. No. 5.5e-31;
Matches 120; Conservative 73; Mismatches 143; Indels 46; Gaps 13;
```

```

QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCSKLMKRLQODKVLITELI--QHSAGHK 76
DB 39 LPSSG-GKCLLDRCRFLAHSAGYILGIVNRCNTIVRR-AGKSVLEQILPAEEVYAR 96
QY 77 VDIDCSQKVVYDQSSQDVASLSDCFVLVLGLKLEKSF--NSVHLAGFAEFSCFPG 134
DB 97 LRSGLYSAVIYDERSPRAESLRDSTVSLVQALRRNAERTDICLKGYERFSSEYBE 156
QY 135 LCEGKSTL-----VPTCISQP-----CLPVA-----NIGFTRILPNLYGCGRDVYNKEL 179
DB 157 FCSKTKALAAIPPPVPSPATEPLDLDCSSCGTPRHDOEGPVEILPFLYLGSAYHAARDM 216
QY 180 IQONGIGVYLAASYCPKPDPIPESHFLRPVNDSPCEKILPWLKSVDFTEKAKASNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQYKCIPEVDNHKADISSWFMELIETIDAVKDCRGR 275
QY 240 VLVHCLAGISRSATLAIYIKRMDSIDEAVRFVYKERTPTSPNPNFLOLDYEKKIK 299
DB 276 VLVHCOAGISRSATLAIYIKRMDSIDEAVRFVYKERTPTSPNPNFLOLDYEKKIK 299
QY 300 NOTGASGPKSKLKLHLEKNEPVAVSEGGQKSETPLSPCADSATSEAAQRPVHPAS 359
DB 336 ATSCAAEAS-----PSGFL-----GERGKTATP-----TSQPFVFSFPV-SVG 373
QY 360 VPSVPSVOPSLIEDSPVQALS 381
DB 374 VHSAPSSLPYL--HSPITTS 393
```

```

RESULT 9
US-09-614-124B-805
; Sequence 805, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fangier, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614.124B
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-614-124B-805
```

```

Query Match          13.3%; Score 453; DB 4; Length 394;
Best Local Similarity 31.4%; Pred. No. 5.5e-31;
Matches 120; Conservative 73; Mismatches 143; Indels 46; Gaps 13;
```

```

QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCSKLMKRLQODKVLITELI--QHSAGHK 76
DB 39 LPSSG-GKCLLDRCRFLAHSAGYILGIVNRCNTIVRR-AGKSVLEQILPAEEVYAR 96
QY 77 VDIDCSQKVVYDQSSQDVASLSDCFVLVLGLKLEKSF--NSVHLAGFAEFSCFPG 134
DB 97 LRSGLYSAVIYDERSPRAESLRDSTVSLVQALRRNAERTDICLKGYERFSSEYBE 156
QY 135 LCEGKSTL-----VPTCISQP-----CLPVA-----NIGFTRILPNLYGCGRDVYNKEL 179
DB 157 FCSKTKALAAIPPPVPSPATEPLDLDCSSCGTPRHDOEGPVEILPFLYLGSAYHAARDM 216
QY 180 IQONGIGVYLAASYCPKPDPIPESHFLRPVNDSPCEKILPWLKSVDFTEKAKASNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQYKCIPEVDNHKADISSWFMELIETIDAVKDCRGR 275
```

```

DB 39 LPSSG-GKCLLDRCRFLAHSAGYILGIVNRCNTIVRR-AGKSVLEQILPAEEVYAR 96
QY 77 VDIDCSQKVVYDQSSQDVASLSDCFVLVLGLKLEKSF--NSVHLAGFAEFSCFPG 134
DB 97 LRSGLYSAVIYDERSPRAESLRDSTVSLVQALRRNAERTDICLKGYERFSSEYBE 156
QY 135 LCEGKSTL-----VPTCISQP-----CLPVA-----NIGFTRILPNLYGCGRDVYNKEL 179
DB 157 FCSKTKALAAIPPPVPSPATEPLDLDCSSCGTPRHDOEGPVEILPFLYLGSAYHAARDM 216
QY 180 IQONGIGVYLAASYCPKPDPIPESHFLRPVNDSPCEKILPWLKSVDFTEKAKASNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQYKCIPEVDNHKADISSWFMELIETIDAVKDCRGR 275
QY 240 VLVHCLAGISRSATLAIYIKRMDSIDEAVRFVYKERTPTSPNPNFLOLDYEKKIK 299
DB 276 VLVHCOAGISRSATLAIYIKRMDSIDEAVRFVYKERTPTSPNPNFLOLDYEKKIK 299
QY 300 NOTGASGPKSKLKLHLEKNEPVAVSEGGQKSETPLSPCADSATSEAAQRPVHPAS 359
DB 336 ATSCAAEAS-----PSGFL-----GERGKTATP-----TSQPFVFSFPV-SVG 373
QY 360 VPSVPSVOPSLIEDSPVQALS 381
DB 374 VHSAPSSLPYL--HSPITTS 393
```

```

RESULT 10
US-09-671-325-805
; Sequence 805, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fangier, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671.325
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-671-325-805
```

```

Query Match          13.3%; Score 453; DB 4; Length 394;
Best Local Similarity 31.4%; Pred. No. 5.5e-31;
Matches 120; Conservative 73; Mismatches 143; Indels 46; Gaps 13;
```

```

QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCSKLMKRLQODKVLITELI--QHSAGHK 76
DB 39 LPSSG-GKCLLDRCRFLAHSAGYILGIVNRCNTIVRR-AGKSVLEQILPAEEVYAR 96
QY 77 VDIDCSQKVVYDQSSQDVASLSDCFVLVLGLKLEKSF--NSVHLAGFAEFSCFPG 134
DB 97 LRSGLYSAVIYDERSPRAESLRDSTVSLVQALRRNAERTDICLKGYERFSSEYBE 156
QY 135 LCEGKSTL-----VPTCISQP-----CLPVA-----NIGFTRILPNLYGCGRDVYNKEL 179
DB 157 FCSKTKALAAIPPPVPSPATEPLDLDCSSCGTPRHDOEGPVEILPFLYLGSAYHAARDM 216
QY 180 IQONGIGVYLAASYCPKPDPIPESHFLRPVNDSPCEKILPWLKSVDFTEKAKASNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQYKCIPEVDNHKADISSWFMELIETIDAVKDCRGR 275
```

QY 240 VLVHCLAGISRSATIAIYIMKMDSLDEAYRFVKEKPTISPNFNLGQLLDYEKKIK 299
 Db 276 VLVHCGIGISRSATICLAYIMMKRVLEBAFEVQKRRIISBPNFSGQLLOFESQVL 335
 QY 300 NOTGASGPKSKLKLHLEKNEPVPVASEGOKSETPSPCADSATSEAAQORPVHPAS 359
 Db 336 ATSCAAMAAAS-----PSGPL-----GERKTPATP-----TSQVFSPFPV-SVG 373
 QY 360 VPSVPSVQPSLLEDSPLVQALS 381
 Db 374 VHSAPSSLPYL--HSPITTSPS 393

RESULT 11
 US-09-589-184-805
 ; Sequence 805, Application US/09589184
 ; Patent No. 6686447
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedvick, Tom
 ; APPLICANT: Carter, Darick
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.478C8
 ; CURRENT APPLICATION NUMBER: US/09/589,184
 ; NUMBER OF SEQ ID NOS: 827
 ; SOFTWARE: FaastSeq for Windows Version 3.0
 ; SEQ ID NO 805
 ; LENGTH: 394
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-589-184-805

Query Match. 13.3%; Score 453; DB 4; Length 394;
 Best Local Similarity 31.4%; Pred. No. 5.5e-31;
 Matches 120; Conservative 73; Mismatches 143; Indels 46; Gaps 13;
 QY 19 LESGTEKVLILDSRPVEYNTSHLEAININCSKLMKRLQODKVLITELI--QHSAMK 76
 Db 39 LPFG-GKCLILDCRPLAHAGYILGSVNVRCNTIVRR-AKGSVLEQILPAEEVRRAR 96
 QY 77 VVDICGQKVVVVOSSQDVASLSDCFLTVLGLKEKSF--NSVHLLAGFAFSPRCFPG 134
 Db 97 LRSGLSAVIVYDERSPRASLREDSTVSLVQALRRNARTRDILKKGGERPSSSEYPE 156
 QY 135 LCEGKSTL-----VPTCISOP-----CLPVA-----NIGPTRLIPNLVLCQORDVINKEL 179
 Db 157 FCGKTKALAIAPPVPSATPEPLDDSCGTPHLHDEGVEILPFLYLGSATHAARDM 216
 QY 180 IQONGIGYVUNASYTCRPFIPESHFLKVPVNDSCFCEKILPWLDSVDFIEKAKSNGC 239
 Db 217 LDALGITALLNVSSDCPN-HFEGHYQKCIPEVDNHRKADISSWMEAIETVIDAVKDCGR 275
 QY 240 VLVHCLAGISRSATIAIYIMKMDSLDEAYRFVKEKPTISPNFNLGQLLDYEKKIK 299
 Db 276 VLVHCGIGISRSATICLAYIMMKRVLEBAFEVQKRRIISBPNFSGQLLOFESQVL 335
 QY 300 NOTGASGPKSKLKLHLEKNEPVPVASEGOKSETPSPCADSATSEAAQORPVHPAS 359
 Db 336 ATSCAAMAAAS-----PSGPL-----GERKTPATP-----TSQVFSPFPV-SVG 373
 QY 360 VPSVPSVQPSLLEDSPLVQALS 381
 Db 374 VHSAPSSLPYL--HSPITTSPS 393

RESULT 12
 US-08-530-290-23
 ; Sequence 23, Application US/08530290
 ; Patent No. 5958721
 ; GENERAL INFORMATION:
 ; APPLICANT: Marshall, Christopher John
 ; APPLICANT: Ashworth, Alan
 ; APPLICANT: Hughes, David Anthony
 ; TITLE OF INVENTION: Methods for Screening of Substances for
 ; TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/530,290
 ; FILING DATE: 14-DEC-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/GB94/00694
 ; FILING DATE: 31-MAR-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9402573.1
 ; FILING DATE: 10-FEB-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9307250.2
 ; FILING DATE: 07-APR-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bastian, Kevin L.
 ; REGISTRATION NUMBER: 34,774
 ; REFERENCE/DOCKET NUMBER: 084611-000000US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 23:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 394 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-08-530-290-23

Query Match. 13.0%; Score 445; DB 2; Length 394;
 Best Local Similarity 31.2%; Pred. No. 2.7e-30;
 Matches 119; Conservative 73; Mismatches 144; Indels 46; Gaps 13;
 QY 19 LESGTEKVLILDSRPVEYNTSHLEAININCSKLMKRLQODKVLITELI--QHSAMK 76
 Db 39 LPFG-GKCLILDCRPLAHAGYILGSVNVRCNTIVRR-AKGSVLEQILPAEEVRRAR 96
 QY 77 VVDICGQKVVVVOSSQDVASLSDCFLTVLGLKEKSF--NSVHLLAGFAFSPRCFPG 134
 Db 97 LRSGLSAVIVYDERSPRASLREDSTVSLVQALRRNARTRDILKKGGERPSSSEYPE 156
 QY 135 LCEGKSTL-----VPTCISOP-----CLPVA-----NIGPTRLIPNLVLCQORDVINKEL 179
 Db 157 FCGKTKALAIAPPVPSATPEPLDDSCGTPHLHDEGVEILPFLYLGSATHAARDM 216
 QY 180 IQONGIGYVUNASYTCRPFIPESHFLKVPVNDSCFCEKILPWLDSVDFIEKAKSNGC 239
 Db 217 LDALGITALLNVSSDCPN-HFEGHYQKCIPEVDNHRKADISSWMEAIETVIDAVKDCGR 275
 QY 240 VLVHCLAGISRSATIAIYIMKMDSLDEAYRFVKEKPTISPNFNLGQLLDYEKKIK 299

Db 276 VLVHCQAGISRSKTTICLAYITMMKKRVLEAEAFYKORSITISPSFSMGQLQFESQVL 335
 QY 300 NQTGASGPKSKLMLHLERKNPEVPVAVSGGQKSETPLSPPCADSATSEAGQRPVPAS 359
 Db 336 ATSCAAEAAAS-----PSGDL-----RRGRTPATP-----ISQFVFSRPV-SVG 373
 QY 360 VPSPVSPQSLLEDSPVLQALS 381
 Db 374 VHSAPSSLPLYL--HSPITTSPS 393

```

RESULT 13
US-09-702-705-827
Sequence 827, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darriick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jame
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702.705
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 827
LENGTH: 394
TYPE: PRT
ORGANISM: Homo sapiens
US-09-702-705-827

```

Query Match	13.0%;	Score 445;	DB 4;	Length 394;
Best Local Similarity	31.4%;	Pred. No. 2.7e-30;		
Matches	120;	Conservative 75;	Mismatches 141;	Indels 46;
				Gaps 14
QY	19	LESGETEYLLIDSRFEVEYNTSHILEATININSKIMKRLQODKYITELI--QHSAXKH	76	
Db	39	LPSSG-GKCLLDRCRFFLAHSAGYILIGSVVNRNTIVRRR-AGAGSVLEBILPAEEVRRAR	96	
QY	77	VDIDCSQKVVYVDQSSQDVASLSSPCFLTVLIGLKEKF--NSVHLLAGFAEFRCFFG	134	
Db	97	LRSLGSAVIYDERSPRAESIREDSVTSLVVOALRPNRNERTDICLKGYERFSSEYDE	156	
QY	135	LCEGKSTL-----VPTGISOP-----C-LFVANI--GPTRIIPNLVYGCGRDVLNREL	179	
Db	157	FCSKTKALAAIPRPYPBSATBELDGCSSCGPRLHDQGGVEIILPPLVYASVHAARDMM	216	
QY	180	IQQNGIIGVYLNASYTCPKRPDIPIRESHFLRPVNDSEFCEKILFWLDSYDFLEKARASNGC	239	
Db	217	LDALGITALLNWSSPDCPN-HFEQGHYQKCIPIYEDNHKADISSMFMEALIEYIDAVDCCGR	275	
QY	240	VLVHCLAIISRSATIAIAYIMKRMMSLDEAATPYKEKRPPTISPNFNELGQLLYEKKIK	299	
Db	276	VLVHCQAIISRSATICLAYIMKKRKRVEEAEEFYKORRSIISPNFSPMGQLQFESQVL	335	
QY	300	NOTASGKSKIKLHLEKPENPVPAVSEGQSKSTPLSPPCADGATSEAGORVHHAS	359	
Db	336	ATSCAAEAS-----PSGPL-----RERKTIPTATP-----TSQVFSPFV-SVG	373	
QY	360	VPSVPSVQPSLLEDSPLVQALS	381	
Db	374	VHSAPSSLPYL--HSPITTPSS	393	

```

US-09-736-457-827
; Sequence 827: Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaltanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvik, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 827
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-827

```

	Query Match	Similarity	Score	DB	Length
Best Local	120	31.4%	445	2.7e-30	394
Matches	120	Conservative	75	Mismatches	141
				Indels	46
				Gaps	14
Qy	19	LESGTEKVLIDSRPVEYNTSHILEAININCSKLMKRRLODDQKYTELI--QHSAAKIK			
Db	39	LEPSG-GKCLLDRCRFLHSAGYILLGSVNRCTIYARR-AKGSVELLEQILPREEVARR			
Qy	77	VVIDDSQKVVYDDSSQVVALSSDCFLTYLLGKLEKSF--NSVHLIAGPAPFSPCPFG			
Db	97	LESGLYSAVIYDEKSPRAESLRERDSIVSLVVALRMRNERTDILCLKGGERFSESYPE			
Qy	135	LCEGKSTL-----VPTGISQP-----C-LPVAANI-GPTRIILPULYVGCQDVIANKEL			
Db	157	FCSKTKALAIIPPIVPEATPEPLDGGSSCQTPHQQGQVVELLPYLYGSAVHAARDM			
Qy	180	IQONGIYVLNASTYCPKPDFIPESHFLRVVPVNDSCFKILPWLDSKVDIFEKAAANGCC			
Db	217	LDALGITALLWVSSDCPN-HEEGHYQKCIPIVDNKNHAKISSWFEAIEYIDAVKDCGR			
Qy	240	VDVHCLAGISRSATIAIAYIMKRMDSLDVAAYFVEKKPPTISPNPNPLGOLLDEVKKIK			
Db	276	VLVHCQAGIISASATICLAIVLMMKRAVLEAFEFVQORSIILSPNSFEGQLAQESQVL			
Qy	300	NOTGASGPKSKLKLHLHEKNEPVPAVSEGGQSETPLSPCCADASTSEAAQORVHAPS			
Db	336	ATSCAAEAAS-----PSGPL-----REKGIPTAPP-----TSGPVFVSFPV-SVG			
Qy	360	VFSVSVSVQPSLLIEDSPVQALIS			
Db	374	VHSAPSSLPYL-HSPITTS			

RESULT 15
US-09-614-124B-827
Sequence 827, Application US/09614124B
Patent No. 6630574
GENERAL INFORMATION:
APPLICANT: Wang, Tongrong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedlick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane

```

1  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
2  TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
3  FILE REFERENCE: 210121.478C9
4  CURRENT APPLICATION NUMBER: US/09/614,124B
5  CURRENT FILING DATE: 2001-07-11
6  NUMBER OF SEQ ID NOS: 1668
7  SOFTWARE: FASTSEQ for Windows Version 3.0
8  SEQ ID NO 827
9  LENGTH: 394
10 TYPE: PRT
11 ORGANISM: Homo sapiens
12 US-09-614-124B-827

```

Query Match	13.0%;	Score 445;	DB 4;	Length 394;	.
Best Local Similarity	31.4%;	Pred. No. 2.7e-30;			
Matches 120;	Conservative 75;	Mismatches 141;	Indels 46;	Gaps 14;	

QY	19	LESGTEKLLIDSRFEVENTSHIIIEAININSKMKRRLOQKVLITTEI--QHSAKHK	76
Db	39	LPSSG--GKCLLIDCRFFLHMSAGIILGSAVVRCTIYRR--AKGSVSLBQLPMBEEVRAR	96
QY	77	VDIDCSQRVVYDQSSQDVASLSSDFCLTVLLGKLEKSF--NSVHLLAGFAEFSRCFPG	134
Db	97	LRSLGYSAVIYEDERSPRAESLRBDSTVSLVVOALRNRMBRTDICLLKGKYEFSSTYPE	156
QY	135	LCEGKSTL-----VPTCISQP-----C-LPVANI--GPTRLPNULYIGCCRDYLNKEL	179
Db	157	FCSKTKALAAIPPPVPPSPATBPLDLGSCSGTPELHQGQPVLELPTLYLGSAYHAARQDM	216
QY	180	IQQNGIGVYLNASVYCCPKEDPFIRESHFLEVPVNDSCFCEKILPMLDVSVPTEKAKASNGC	239
Db	217	LDALGITLALNVSSDCPN--HFEGHYQYKCI PVEDNHKADISSWFMEAIETIIDVAKCGRGR	275
QY	240	VLVHGLAGISRSATIAIAYIMKRMDSLDEAARYFVEKEKPTISPNENFLGQLLDYEEKIX	299
Db	276	VLVHGOAGISRSATICTIAYLMMKKRVRLDEAFEFVQORSIISPNNSFMGQLLQFSSQYL	335
QY	300	NOTGASGPKSKKLLHLEKPYNEVPVAVSSGQKSETPPLSPCADSAITSEAAQGRPVHPMS	359
Db	336	ATSCAAEAS-----PSGPLY-----REKGTBPATP-----TSQFVSFPV--SYG	373
QY	360	VPSVPVPOSLLEDSPVQALS	381
Db	374	VHSAPDSLPYL--HSPITTSPS	393

Search completed: June 21, 2004, 13:25:43
Job time : 23.6939 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:15:52 ; Search time 10.3061 Seconds
(without alignments)
2818.703 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302

Sequence: 1 MAHEMICTQVTERVALLE.....PNRNFGLQGLDYERKIKNQK 302

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	450	28.8	619	2	T15969
2	425	27.2	367	1	S24411
3	423	27.0	394	2	A56115
4	418	26.7	314	1	A57126
5	415	26.5	367	1	S29090
6	415	26.5	367	2	S52265
7	409	26.2	314	2	B57126
8	377.5	24.1	393	2	A56947
9	351	22.4	384	1	T13890
10	319	20.4	303	2	T46405
11	289	18.5	365	2	T32494
12	223.5	14.3	330	2	T39698
13	222.5	14.2	186	2	T16056
14	212.5	13.6	226	2	T21380
15	208	13.3	272	2	T18915
16	207.5	13.3	220	2	JC7885
17	205	13.1	223	2	T49365
18	202	12.9	364	1	S31304
19	201.5	12.9	142	2	T03074
20	195.5	12.5	185	1	A47126
21	195	12.5	283	2	G84458
22	193	12.3	205	2	T49364
23	185	11.8	807	1	S44538
24	183.5	11.7	276	2	T48906
25	181.5	11.6	278	2	T39517
26	179.5	11.5	866	2	F88481
27	173.5	11.1	204	2	T17802
28	171	10.9	489	1	S58725
29	162	10.4	771	2	T47666

30	160.5	10.3	580	2	T18439	hypothetical prote
31	160.5	10.3	600	2	T18446	hypothetical prote
32	157	10.0	209	1	S48459	probable dual spec
33	152	9.7	282	2	S41012	hypothetical prote
34	150.5	9.6	169	2	T30684	probable dual spec
35	147.5	9.4	597	1	S43743	probable dual spec
36	145.5	9.3	171	1	T36845	dual specificity p
37	145.5	9.3	171	2	T28522	probable dual spec
38	145.5	9.3	171	2	B72161	probable dual spec
39	144.5	9.2	171	1	Q0V2H1	dual specificity p
40	138.5	8.9	171	1	A42514	dual specificity p
41	134	8.6	150	2	T21489	hypothetical prote
42	134	8.6	272	2	T19418	hypothetical prote
43	127.5	8.2	171	1	B47452	dual specificity p
44	121	7.7	928	1	S50578	hypothetical prote
45	117.5	7.5	160	2	T10278	protein tyrosine p

ALIGNMENTS

```
RESULT 1
T15969
hypothetical protein F08B1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15969
R:Chisoe, S.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of C. elegans cosmid F08B1.
A:Reference number: Z18439
A:Accession: T15969
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-619 <CHI>
A:Cross-references: EMBL:U23178; NID:9726421; PID:9726422; PIDN:AAC46719.1; CESP:F08B1.1
A:Experimental source: strain Bristol N2
C:Genetic:
A:Gene: CESP:F08B1.1
A:Introns: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3

Query Match          28.8%; Score 450; DB 2; Length 619;
Best Local Similarity 50.6%; Pred. No. 2.3e-29;
Matches 90; Conservative 34; Mismatches 48; Indels 6; Gaps 4;

QY 123 GGAPEFRCPGLCEGKS--TLVPTCTSPCL--PVANIGPTRILPVLVYGCQPDVANK 178
      |||:::|
Db 99 GGFQFAQYQPOLCESESGWTRLPQSLQPCLSQPTD-GITLTPNIVYLSQIDSDT 157
QY 179 LIQONGIGVYVNASYTPKPKDFPE-SHFLRVPNVDFCEKILPWLKSVDFEKASVN 237
      ::|||:::|
Db 158 MLALDLSVINISMTCPKVCIKEDNFMRIPIVNDYSQKLSYFPMAVFEKCRAG 217
QY 238 GCVLVHCLAGISRSATIAIAYIMKRMDSLEAVRFVKEKRPITSPNFNFGQLDYE 295
      |||:::|
Db 218 KCLIHCLAGISRSPTIAISYIMRYMKGSDDAIVRYKERPSISPNFNFQGLDYE 275

RESULT 2
S24411
dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - mouse
N:Alternate names: 3CH134 protein; protein-tyrosine-phosphatase erp, nonreceptor type 10
C:Species: Mus musculus (house mouse)
C:Date: 19-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999
C:Accession: A54681; S24411
R:Noguchi, T.; Metz, R.; Chen, L.; Mattei, M.G.; Carrasco, D.; Bravo, R.
Mol. Cell. Biol. 13, 5195-5205, 1993
A:Title: Structure, mapping, and expression of erp, a growth factor-inducible gene encod
A:Reference number: A54681; MIM:3360956; PMID:8355678
A:Accession: A54681
A:Molecule type: DNA
A:Residues: 1-367 <NOG>
A:Cross-references: GB:S64851; NID:9409976; PIDN:AAB27882.1; PID:9409977
```

R,Charles, C.H.; Abler, A.S.; Lau, L.F.
Oncogene 7, 187-190, 1992
A:Title: cDNA sequence of a growth factor-inducible immediate early gene and characteriz
A:Reference number: S24411; MUID:92158357; PMID:1741163
A:Accession: S24411
A:Molecule type: mRNA
A:Residues: 1-367 <CHA>
A:Cross-references: EMBL:X61940; NID:949735; PID:CAA43944.1; PID:949736
A:Genetics: erp
A:Introns: 123/1; 172/1; 245/1
C:Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity
C:Keywords: immediate-early protein; phosphoprotein; phosphoric monoester hydrolase
F:181-312/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
F:256/Active site: Cys (phosphocysteine intermediate) #status predicted
F:264/Binding site: substrate phosphate (Arg) #status predicted

Query Match
Best Local Similarity 27.2%; Score 425; DB 1; Length 367;
Matches 104; Conservative 48; Mismatches 129; Indels 22; Gaps 5;

QY 15 LVALLSGTEKVLIDSRPFVEYNTSHILEAININCKLMKRLQODKYLITELIOHSAR 74
Db 13 LRALLREGAQCCLLDCRSFPFAMNCHINGVAVRFTTVRRRAKMGLEHIVPAELR 72
QY 75 HKVDIDCSQKVVVYDSSQDVASLSDCFVLVLGKLT--EKSFNVSVHLAAGFAEFRCF 132
Db 73 GRLLAGVAVVLDLDRSASLDGAKRDTGLALAGALCREARSTQVFLQGVHARSASC 132
QY 133 PGLCEKSTLVPTCTISQP-----CLPVANI-----GPTRLIPNLVYGCQRDYL 175
Db 133 PELCKQKST--PTGLSLPLSTSVDPBSAEGSSGCTPLVYDGGVELLPLVYLSAYVAS 190
QY 176 NKEILIQNGIGVNLVASYTCRPPDFIESHFLRVVNDSCFCKILPMLDKSVDFIKAKA 235
Db 191 RQMDLALGITALINVSANCPN-HFECHYQYKSLPVEDNHRKADISSWFEAIDFIDSTID 249
QY 236 SNGCVLVHCLAGISRSATLAIYIMKMDSLDEAVRFVKEKPTISPNFNLGQLDYE 295
Db 250 AGGRVFNVCQAGISRSATICLAIYLMRNVKLEDAPEFVQKRSITISPNFSGQLDQFE 309
QY 296 KKI 298
Db 310 SQV 312

RESULT 3

A56115
dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 4 - human
N:Alternate names: dual specificity phosphatase HVH2
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 24-Apr-1998
C:Accession: A56115
R:Guam, K.L.; Butch, R.
J. Biol. Chem. 270, 7197-7203, 1995
A:Title: Isolation and characterization of a novel dual specific phosphatase, HVH2, whic
A:Accession: A56115
A:Reference number: A56115; MUID:95221370; PMID:7555768
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-394 <GUA>
A:Cross-references: GB:U21108
C:Genetics:
A:Gene: GDB:DUSP4; HVH2; MKP-2
A:Cross-references: GDB:433893
A:Map position: 8p21-8p11.2
C:Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity
C:Keywords: phosphoprotein; phosphoric monoester hydrolase
F:203-334/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
F:286/Active site: Cys (phosphocysteine intermediate) #status predicted
F:286/Binding site: substrate phosphate (Arg) #status predicted

Query Match
27.0%; Score 423; DB 2; Length 394;

Best Local Similarity 33.4%; Pred. No. 2.2e-27;
Matches 100; Conservative 63; Mismatches 114; Indels 22; Gaps 9;

QY 19 LBSGTEKVLIDSRPFVEYNTSHILEAININCKLMKRLQODKYLITELI--OHSARK 76
Db 39 LPBG-GKCLLDRCPLAHSAGVILGSVAVRNTIVRR-ANGSVSLQIIPAEVEVRAR 96
QY 77 VDIDCSQKVVVYDSSQDVASLSDCFVLVLGKLEKSF--NSVHLAAGFAEFRCFPG 134
Db 97 LRSGLSAVIYVDSSPRAESLRDSTVSLVQALRRNAERTDICLKGVERFSSEYFE 156
QY 135 LBSGKSTL-----VPTCTISQP-----C-LPVANI--GPTRLIPNLVYGCQRDYL 179
Db 157 FCSKTKALAAIPPPVPPEATPEPLDGCSSGCTPLHDGCGVELLPLVYLSAYVAARRDM 216
QY 180 IQONGIGVNLVASYTCRPPDFIESHFLRVVNDSCFCKILPMLDKSVDFIKAKASNGC 239
Db 217 LDALGITLALNVSSDCRN-HFECHYQYKSLPVEDNHRKADISSWFEAIEYIDAVDCRRR 275
QY 240 VLVHCLAGISRSATLAIYIMKMDSLDEAVRFVKEKPTISPNFNLGQLDYEKKI 298
Db 276 VLVHCLAGISRSATICLAIYLMRNVKLEDAPEFVQKRSITISPNFSGQLDQFE 334

RESULT 4

A57126
dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 2 - human
N:Alternate names: mitogen-induced nuclear tyrosine phosphatase; phosphatase of activat
C:Species: Homo sapiens (man)
C>Date: 03-Nov-1995 #sequence_revision 02-Jul-1996 #text_change 11-Jun-1999
C:Accession: A57126
R:Roman, P.J.; Davis, P.; Moskaluk, C.A.; Keane, M.; Krutzsch, H.; Siebenlist, U.; Keil
Science 259, 1763-1766, 1993
A:Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
A:Reference number: A57126; MUID:93206122; PMID:7681221
A:Accession: A57126
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-314 <ROH>
A:Cross-references: GB:L11329; NID:9559539; PID:AAA50779.1; PID:9292376
C:Genetics:
A:Gene: GDB:DUSP2
A:Cross-references: GDB:139200
A:Map position: 2q11-2q11
C:Function:
A:Description: catalyzes the hydrolysis of peptidyl-phosphoserine, -phosphothreonine, an
C:Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity
C:Keywords: nucleus; phosphoprotein; phosphoric monoester hydrolase
F:140-160/Region: nuclear location signal
F:180-311/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
F:257/Active site: Cys (phosphocysteine intermediate) #status predicted
F:263/Binding site: substrate phosphate (Arg) #status predicted

Query Match
Best Local Similarity 26.7%; Score 418; DB 1; Length 314;
Matches 109; Conservative 45; Mismatches 109; Indels 44; Gaps 8;

QY 24 EKVLLIDSRPFVEYNTSHILEAININCKLMKRR-----LOQDKYLITELIOHS 72
Db 25 ERTLLIDCRPLAFRCRHVRAARVPVNALIRRRARRGPPAVLACLPDRALRRLVARGE 84
QY 73 AKHKVDIDCSQKVVVYDSSQDVASLSD-----CFLVLVLGKLEKSFNSVHLAAGFAEF 128
Db 85 L-----AAAVLDEGSASVALRDPSPAHVLAALHETRAGTAVYFRLGGDFGR 135
QY 129 SBCFPGLC-EKGSSTLV-----TCISQPCLPVANI--GPTRLIPNLVYGCQRDYLANKEL 181
Db 136 QCCCPDLCSBAPALPPTGKTRSDSRAPVVDGGPVILTYLFLGSCSHSSDQGG 195
QY 182 QNGIGVNLVASYTCRPPDFIESHF-----LRPVNDSCFCKILPMLDKSVDFIEKAKA 235
Db 196 ACGITAVLVASACP-----NHFEGLFRYKSIIPVEDNQMVEISMPQEARLIGFTIDWVKN 248

QY 236 SNGCVLVHCLAGISRSATIAIAYIMKMDVSLDEAYRFEVKEKPTISPNFPGQLLDYE 295
 DB 249 SGRRLVHCAQGISRSATITCLATLMQSRRLDEAFVQKRGVISPNSFMGQLLOFE 308
 QY 296 KK1 298
 DB 309 TQV 311

RESULT 5

S29090
 dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - human
 N/Alternate names: protein-tyrosine-phosphatase C1100; protein-tyrosine-phosphatase, not
 C/Species: Homo sapiens (man)
 C/Date: 25-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999
 C/Accession: S29090; A53052
 R/Keyes, S.M.; Emalle, E.A.
 Nature 359, 644-647, 1992
 A/Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine
 A/Reference number: S29090; MUID:93024952; PMID:1406996
 A/Accession: S29090
 A/Molecule type: mRNA
 A/Residues: 1-367 <KEY>
 A/Cross-references: EMBL:X68277; NID:G29980; PIDN:CAA48338.1; PID:G29981
 R/Kwak, S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.
 J. Biol. Chem. 269, 3596-3604, 1994
 A/Title: Isolation and characterization of a human dual specificity protein-tyrosine pho
 A/Reference number: A53052; MUID:94148864; PMID:8106404
 A/Accession: A53052
 A/Molecule type: DNA
 A/Residues: 1-367 <KWA>
 A/Experimental source: leukocyte
 A/Note: sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804,
 C/Genetics:
 A/Gene: GDB:DUSP1; PTPN10
 A/Cross-references: GDB:136197; OMIM:600714
 A/Map position: 5q34-5q34
 C/Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity
 C/Keywords: heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induced P
 F/181-312/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>
 F/258/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 26.5%; Score 415; DB 1; Length 367;
 Best Local Similarity 33.9%; Pred. No. 9.2e-27;
 Matches 102; Conservative 47; Mismatches 134; Indels 18; Gaps 4;

QY 15 LVALLSGTEKVLIDSRPVEYNTSHILEAININCSKMKRRLOQDKVLTITLIGHSK 74
 DB 13 LRLALLGERAAQCILLDCRSFFAFNAGHAGSVNVRSTIVRRRAKMGMEHIVPNAELR 72
 QY 75 HKYVDICDSQKVVYVYDSSQDVASLSDCFLTVLIGKL--EKSFNVSVHLLAGFAEFGRCP 132
 DB 73 GRLLAGVHVAHVLLDERSALDGAKRDDGTALLAAGALCREARAQVFLGGYEAFASGC 132
 QY 133 PGICBEGK-----STVPTCISQCLPVANI-----GPRILPMLYLGCGRDVLNK 177
 DB 133 PELCSKQSTPMGSLPLSTVPSDAESGSSCSPTLYDQGPVELLPFLVLSAYHNASRK 192
 QY 178 ELIQONGIGVVLNASTYCPKDPFIPESHFLRPVNDSCFKILPWLDSVDFIEKAKASN 237
 DB 193 DMDALGITLINVANSANCPN-HFEGHYQKSIPEVDNKKADISSWFEMALDFIDSTINAG 251
 QY 238 GCVLVHCLAGISSATIAIAYIMKMDVSLDEAYRFEVKEKPTISPNFPGQLLDYEKK 297
 DB 252 GRVFEVHCQAGISSATITCLAVIMKTRNVKLDFAEFVQKRSITSPNSFMGQLLOFE 311
 QY 298 I 298
 DB 312 V 312

RESULT 6

S52265
 dual specificity phosphatase (EC 3.1.3.-) 1 - rat
 N/Alternate names: protein-tyrosine-phosphatase C1100; protein-tyrosine-phosphatase, noni
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 22-Jun-1999
 C/Accession: S52265
 R/Muda, M.; Schlegel, W.; Arkinstall, S.
 submitted to the EMBL Data Library, January 1995
 A/Description: Pathways regulating C1100 gene expression in pituitary cells.
 A/Reference number: S52265
 A/Accession: S52265
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-367 <KUD>
 A/Cross-references: EMBL:X64004; NID:G642264; PIDN:CAA58828.1; PID:G642265
 C/Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity
 C/Keywords: phosphoprotein; phosphoric monoester hydrolase
 F/181-312/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>
 F/258/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 26.5%; Score 415; DB 2; Length 367;
 Best Local Similarity 34.0%; Pred. No. 9.2e-27;
 Matches 104; Conservative 48; Mismatches 126; Indels 28; Gaps 6;

QY 15 LVALLSGTEKVLIDSRPVEYNTSHILEAININCSKMKRRLO-----QDKVLTITELI 69
 DB 13 LRLALLGERAAQCILLDCRSFFAFNAGHAGSVNVRSTIVRRRAKMGMEHIVPNTLE- 71
 QY 70 QHSKHKVDICDSQKVVYVYDSSQDVASLSDCFLTVLIGKL--EKSFNVSVHLLAGFAE 127
 DB 72 ----RRLLAGVHVAHVLLDERSALDGAKRDDGTALLAAGALCREARSTGVFLGGYE 127
 QY 128 FSCFPGLCEBK-----STVPTCISQCLPVANI-----GPRILPMLYLGCGR 172
 DB 128 FSASCPCLCSKQSTPMGSLPLSTVPSDAESGSSCSPTLYDQGPVELLSFLYGSAY 187
 QY 173 DVNLKELIQONGIGVVLNASTYCPKDPFIPESHFLRPVNDSCFKILPWLDSVDFIEK 232
 DB 188 HSRKQMLDALGITLINVANSANCPN-HFEGHYQKSIPEVDNKKADISSWFEMALDFID 246
 QY 233 AKASNCVLVHCLAGISRSATIAIAYIMKMDVSLDEAYRFEVKEKPTISPNFPGQL 292
 DB 247 IKDAGRVFPHCAQGISRSATITCLAVIMKTRNVKLDFAEFVQKRSITSPNSFMGQL 306
 QY 293 DYKKI 298
 DB 307 QFESQV 312

RESULT 7

B57126
 dual specificity phosphatase (EC 3.1.3.-) 2 - mouse
 N/Alternate names: mitogen-induced nuclear protein-tyrosine-phosphatase PAC-1
 C/Species: Mus musculus (house mouse)
 C/Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 24-Apr-1998
 C/Accession: B57126
 R/Rohan, P.J.; Davis, P.; Moskaluk, C.A.; Kearns, M.; Kruttsch, H.; Siebenlist, U.; Kelly
 Science 259, 1763-1766, 1993
 A/Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
 A/Reference number: A57126; MUID:93206122; PMID:7681221
 A/Accession: B57126
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-314 <ROH>
 A/Cross-references: GB:111330
 C/Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity
 C/Keywords: nucleus; phosphoprotein; phosphoric monoester hydrolase
 F/180-311/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>
 F/257/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/263/Binding site: substrate phosphate (Arg) #status predicted

Query Match 26.2%; Score 409; DB 2; Length 314;

Best Local Similarity 36.0%; Pred. No. 2,4e-26;
Matches 107; Conservative 45; Mismatches 113; Indels 32; Gaps 7;

```
QY 24 EKVLLIDSRPVEYNTSHILEAININCKLMKRR-----LOODKVLITELIQHS 72
Db 25 EKTLLIDCRPFIAFGRSHVRAARPVPMALLRRRPGTPAALACLIPDALTRKLRGE 84
QY 73 AKHKVDIDCSQKVYVYDSSQDVASLSDCFLVTLGKLEKSF---NSVHLIAGRAEF 128
Db 85 L-----ARAVLDBSSASVTELPDPGPAHLLALQHEMRGPTTCFLRGFKFSF 135
QY 129 SRCPGLC-EKSKTLVP-----TCISQPLCPVANI-GPTRIPLVLYGQCDVNLKELIQ 181
Db 136 QTYCPDLCSFAPALPAPGAENSNDRPVLYDQGGFVELLPVLYIGSCNHSDDQGLQ 195
QY 182 QNGIGVYLNASTYCPKDPFIPESHFLAVPVNDSCFKILPWLDKSDVFIKAVANGCYL 241
Db 196 ACGITAVLVNASTCPN-HFEGLFHYKSIPEVDNQWVAISAMPQDAISFIDSVKNGSRVL 254
QY 242 VHLIAGISRSATTAIAYIMKMDSLDEAVRFVYKERTPTSPNPNFLGQLDYEKKI 298
Db 255 VHCQAGISRSATICTAVLIQSHRRLDEADPFVKQRGVISPVFSFGQLQLETV 311
```

RESULT 8

dual specificity phosphatase (EC 3.1.3.-) HVH2 - rat
N/Alternate names: mtogen-activated protein kinase phosphatase 2
C/Species: Rattus norvegicus (Norway rat)
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 17-Mar-1999
C/Accession: A56947
R/Miscr-Press, A.; Rim, C.S.; Yao, H.; Robertson, M.S.; Stork, P.J.S.
J. Biol. Chem. 270, 14587-14596, 1995
A/Title: A novel mtogen-activated protein kinase phosphatase. Structure, expression, and
A/Reference number: A56947; MUID:95301550; PMID:7782322
A/Accession: A56947
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-393 <WIS>
C/Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity
C/Keywords: phosphoprotein, phosphoric monoester hydrolase
F/202-333/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
F/279/Active site: Cys (phosphocysteine intermediate) #status predicted
F/285/Binding site: substrate phosphate (arg) #status predicted

Query Match 24.1%; Score 377.5; DB 2; Length 393;
Best Local Similarity 31.0%; Pred. No. 1.3e-23;
Matches 95; Conservative 54; Mismatches 124; Indels 33; Gaps 8;

```
QY 16 VALLESGETKVLIDSRPVEYNTSHILEAININCKLMKRLQODKVLITELIQSAKH 75
Db 38 IGLLSGG--KCLLDRCRFLAHSAQYINGSVNRCNTIVRR--AKGSVSLQQLI--PAEE 92
QY 76 KV-----DIDCSQKVYVYDSSQDVASLSDCFLVTLGKLEKSFNSVHLIAGRAEFS 129
Db 93 EVRPCALASTLSSTMTQPARRESPGQHSVAGRALRRNAET--DICLLKGYYRRS 150
QY 130 RCFPLCGSKSTVPTCISQPLCPVANI-----GPTRIPLVLYGQCD 172
Db 151 SEYEPFGCKTAL--AALPPVPSTNESLDLGGSCGTPLDQGGFVELLPVLYIGSAV 208
QY 173 DLNKEILLQONGIGVYNASTYCPKDPFIPESHFLAVPVNDSCFKILPWLDKSDVFIK 232
Db 209 HAARDMDLALGITALNVSDCN-HFEGHYQKCIPEVDNKADISFMEMALTEYIDA 267
QY 233 AKANGCVLVHCLAGISRSATTAIAYIMKMDSLDEAVRFVYKERTPTSPNPNFLGQL 292
Db 268 VKDQKGVLVHCLAGISRSATICTAVLMKMDSLDEAVRFVYKERTPTSPNPNFLGQL 327
QY 293 DYKKI 298
Db 328 QFESQV 333
```

RESULT 9

138890

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 5 - human
N/Alternate names: dual specificity phosphatase B23; dual-specificity phosphatase hvh-3,
C/Species: Homo sapiens (man)
C/Date: 16-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 11-Jun-1999
C/Accession: 138890; A55313
J. Kwak, S.P.; Dixon, J.E.
J. Biol. Chem. 270, 1156-1160, 1995
A/Title: Multiple dual specificity protein tyrosine phosphatases are expressed and regulat
A/Reference number: A55432; MUID:95138103; PMID:7836374
A/Accession: 138890
A/Molecule type: mRNA
A/Residues: 1-384 <RES>
A/Cross-references: EMBL:U6966; NID:9642012; PIDD:AMB06261.1; PID:9642013
A/Experimental source: placenta
R/Ishibashi, T.; Bottaro, D.P.; Micheli, P.; Kelley, C.A.; Aaronson, S.A.
U. Biol. Chem. 269, 29897-29902, 1994
A/Title: A novel dual specificity phosphatase induced by serum stimulation and heat shock
A/Reference number: A55313; MUID:95050849; PMID:7961985
A/Accession: A55313
A/Molecule type: mRNA
A/Residues: 1-8, 'GHV', 12-70, 'R', 72-104, 'F', 107-362, 'RCLPTQSSGSAELMORPNPAKTGMESAPQERQ',
A/Cross-references: GB:U15932; NID:9606971; PIDD:AAA64693.1; PID:9606972
A/Experimental source: mammary epithelial cells
C/Genetics:
A/Gene: GDB:DUSP5
A/Cross-references: GDB:385447
A/Map position: 10q25-10q26
C/Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity
C/Keywords: phosphoprotein, phosphoric monoester hydrolase
F/186-317/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
F/263/Active site: Cys (phosphocysteine intermediate) #status predicted
F/269/Binding site: substrate phosphate (arg) #status predicted

Query Match 22.4%; Score 351; DB 1; Length 384;
Best Local Similarity 29.4%; Pred. No. 2e-21;
Matches 94; Conservative 54; Mismatches 124; Indels 48; Gaps 7;

```
QY 18 LLESGETKVLIDSRPVEYNTSHILEAININCKLMKRLQODKVLITELIQSAKHV 77
Db 15 LKREAAARCVLDGRPYLAFAASNVRSGLNVNVSIVLRARCAV-----SARYVL 66
QY 78 DIDCSQ-----KVYVYDSSQDVASLSDCFLVTLGKLEKSFNS--VHLIA 122
Db 67 PDEAARLILDEGGGVAAVVLDDGSRHMQKLEESAARVLTSLACLIPAGRYVFLK 126
QY 123 GGFARFRCFPLC-----EGSKTLVPTCISQPLCPVANI-----GPTRI 162
Db 127 GGYETFFSYBECCVDVYPIQKIESRALLISQCK---PVNVSVRYPAVDQGGFVEI 182
QY 163 LPNLTLCQCDVNLKELIQONGIGVYNASTYCPKDPFIPESHFLAVPVNDSCFKILPW 222
Db 183 LPFLYIGSAVHASKCEFLANLITLVNSRRTSEA-CMTHLHYKWIPEVDSTHAISSH 241
QY 223 LDKSDVFIKAVANGCVLVHCLAGISRSATTAIAYIMKMDSLDEAVRFVYKERTPT 282
Db 242 FOEALIDPDCVREKGVLVHCLAGISRSPTICMAIYMKTPKQRLKEAFDYIKORSSWS 301
QY 283 PNFNPLQQLDYEKKINQOT 302
Db 302 PNFNPLQQLDYEKKINQOT 321
```

RESULT 10

T46405

hypothetical protein DKFZP434O1321.1 - human
C/Species: Homo sapiens (man)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 21-Jun-2002
C/Accession: T46405
R./Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, U.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000

A:Reference number: 223034
 A:Accession: T46405
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-303 <AAA>
 A:Cross-references: EMBL:AL137704
 A:Experimental source: adult testis; clone DKFZp43401321
 C:Genetics:
 A:Note: DKFZp43401321.1

Query Match 20.4%; Score 319; DB 2; Length 303;
 Best Local Similarity 37.7%; Pred. No. 6.5e-19;
 Matches 72; Conservative 33; Mismatches 70; Indels 16; Gaps 5;

QY 123 GGFARFRCRPGCGEGSTL-----VPTCSQP-----C-LPVANI-GPTRIPLNLY 167
 DB 54 GGYRFRSEYEPFSCSKTKALAIAPPVPSPATEPLDGCSCGCPHLDGQGPVILPFLY 113
 QY 168 LGCORVNLKELIQONGIGVYNASVTCPPKDPFIPESHFLRPVNDSECEKILPMLDKSV 227
 DB 114 LGSAYHARRDMDALGITLALNVSSDCPN-HFEGHQYKCIPEVDNKHADISSWEMAI 172
 QY 228 DFLERKASNGCVLVHCLAGISRSATTATAYIMKMDMSIDEAYRKYKRPITSPNPF 287
 DB 173 EYIDAVKDCRGVLVHCOAGISRSATICLAYIMKKKVRLEAFEPYKORRSIISPNFSF 232
 QY 288 LGQLLDYEK 298
 DB 233 MGQLQFESQV 243

RESULT 11

T32494
 hypothetical protein C05B10.1 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C/Accession: T32494
 R:Geisler, C.; Mameley, P.
 submitted to the EMBL Data Library, December 1997
 A:Description: The sequence of C. elegans cosmid C05B10.
 A:Reference number: 221178
 A:Accession: T32494
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-365 <GEI>
 A:Cross-references: EMBL:AF036685; PIDN:AA88308.1; GSPDB:GN00022; CESP:C05B10.1
 A:Experimental source: strain Bristol N2; clone C05B10
 C:Genetics:
 A:Gene: CESP:C05B10.1
 A:Map position: 4
 A:Introns: 22/2; 46/3; 163/2; 204/2; 248/1; 287/2

Query Match 18.5%; Score 289; DB 2; Length 365;
 Best Local Similarity 24.9%; Pred. No. 2.5e-16;
 Matches 80; Conservative 67; Mismatches 134; Indels 40; Gaps 7;

QY 7 GTQVTERLVALLSEGEKVLIDSRPFVEYNTSHIEAININCSKLMKRLQODKVLIT 66
 DB 12 GBEISAQFNRKIFHE--RNVIVLDCRS---NGDSVRAVRRLRLPALIQRLMGSGWRLS 65
 QY 67 ELIQHSAKHVYDIDCSQKVVVYDOSD-----VASLSDCGLVTLGKLEKSPSVH 119
 DB 66 TVPDLKDLNLSPPDCPEVLILPGDSEDBQSALANLNSNRHFLVLSPVYTLISQF 125
 QY 120 LLAGGFAEFS-----RCFPGLCGKSTLVPTCISQPCLVANIG----- 158
 DB 126 PTLRADADEMWNTPFGNNSMPGASGQ-----ASSGPELNLNLRLLEGDQSGKQAEF 180
 QY 159 PTRLPLNLYGCGQDVNLKELIQONGIGVYNASVTCPPK-DPIESHFLRPVNDSECE 217
 DB 181 PVKLTNPLVLYGNMETAKNRDVLTKKYSISHVINTVSNLPTFEEDPNKRYRISADDNASH 240
 QY 218 KILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSATTATAYIMKMDMSIDEAYRKYK 277

DB 241 NLTKPFPEAISPIDDARRNDSACLVHCLAGISRSVITCLAYIMKTEKCTDSANVSWQKR 300
 QY 278 RPTISPNFNLGQLDYEK 298
 DB 301 NASIAPIFHFHMGQLTDYEKML 321

RESULT 12

T39698
 protein tyrosine phosphatase - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C/Accession: T39698
 R:Wood, V.; Skelton, J.; Churcher, C.M.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, July 1999
 A:Reference number: 221870
 A:Accession: T39698
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-330 <MOO>
 A:Cross-references: EMBL:AL109652; PIDN:CAB51765.1; GSPDB:GN00067
 A:Experimental source: strain 972h-; cosmid c17A3
 C:Genetics:
 A:Gene: p1041
 A:Map position: 2

Query Match 14.3%; Score 223.5; DB 2; Length 330;
 Best Local Similarity 39.3%; Pred. No. 5.9e-11;
 Matches 53; Conservative 18; Mismatches 63; Indels 1; Gaps 1;

QY 160 TRLPLNLYGCGQDVNLKELIQONGIGVYNASVTCPPKDPFIPESHFLRPVNDSECEK 219
 DB 48 SEISKLYISWKTASELVSIPKIDYTLISAMSNPNLS-VBQOHLWLQIEDSSQNT 106
 QY 220 LPMLDKSVDFIEKAKASNGCVLVHCLAGISRSATTATAYIMKMDMSIDEAYRKYKRP 279
 DB 107 LQYFEKSNKRTIARLSKNMAYLVHCFRAGISRSVTLVAAYIMKNNMTTEALSHINERS 166
 QY 280 TISPENFLGQLD 294
 DB 167 GISPNANFLRLQLRVY 181

RESULT 13

T16056
 hypothetical protein F13D11.3 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001
 C/Accession: T16056
 R:Fullon, L.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid F13D11.
 A:Reference number: S69020
 A:Accession: T16056
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-186 <FUL>
 A:Cross-references: EMBL:U40939; NID:g1072175; PID:g1072176; PIDN:AAA81700.1; CESP:F13D11
 C:Genetics:
 A:Gene: CESP:F13D11.3
 A:Introns: 30/3; 57/3; 85/3; 125/3; 172/3

Query Match 14.2%; Score 222.5; DB 2; Length 186;
 Best Local Similarity 36.2%; Pred. No. 3.4e-11;
 Matches 51; Conservative 31; Mismatches 50; Indels 9; Gaps 3;

QY 160 TRLPLNLYGCGQDVNLKELIQONGIGVYNASVTCPPKDPFIPESHFLRPVNDSECE 216
 DB 12 TVRPHFLIAGYGC-----ITPSLKQYNITHGVDCTNLKTR--IKGLRIEVPVDNLT 65
 QY 217 EKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSATTATAYIMKMDMSIDEAYRKYK 276

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:13:37 ; Search time 6.55843 Seconds

(without alignments)
2397.707 Million cell updates/sec

Title: US-10-029-345a-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMTGTQVTERIVALLLE.....PNFNPLGQJLDYERKIKNQRT 302

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1552	99.2	665	1 DUSG_HUMAN	Q9BY84 homo sapien
2	988	63.2	625	1 DUS8_HUMAN	Q13202 homo sapien
3	985	63.0	663	1 DUS8_MOUSE	O09112 mus musculu
4	469	30.0	482	1 DUS8_HUMAN	O9Y666 homo sapien
5	468	29.9	483	1 DUS8_MOUSE	O9E880 mus musculu
6	467.5	29.9	381	1 DUS6_HUMAN	Q16828 homo sapien
7	462.5	29.6	381	1 DUS6_MOUSE	Q9dbb1 mus musculu
8	461.5	29.5	381	1 DUS6_MOUSE	O64346 rattus norv
9	450	28.8	619	1 VHP1_CAEEL	O10038 caenorhabd
10	431	27.6	375	1 DUS4_CHICK	O9Pw71 gallus gall
11	425	27.2	367	1 DUS1_MOUSE	P28563 mus musculu
12	425	27.2	394	1 DUS4_HUMAN	Q13115 homo sapien
13	423.5	27.1	395	1 DUS4_RAT	O62767 rattus norv
14	418	26.7	314	1 DUS2_HUMAN	O05923 homo sapien
15	415	26.5	367	1 DUS1_HUMAN	P28563 homo sapien
16	415	26.5	367	1 DUS1_RAT	O64623 rattus norv
17	409	26.2	318	1 DUS2_MOUSE	O05922 mus musculu
18	408	26.1	384	1 DUS9_HUMAN	O99956 homo sapien
19	385.5	24.6	320	1 DUS7_HUMAN	Q16829 homo sapien
20	382.5	24.5	320	1 DUS7_MOUSE	O91246 mus musculu
21	358.5	22.9	384	1 DUS5_MOUSE	O64838 rattus norv
22	357.5	22.9	280	1 DUS7_RAT	O63340 rattus norv
23	351	22.4	384	1 DUS5_HUMAN	O16690 homo sapien
24	262	16.8	198	1 DUSE_HUMAN	O95147 homo sapien
25	261	16.7	198	1 DUSE_MOUSE	O9j117 mus musculu
26	205.5	13.1	188	1 DUSJ_HUMAN	O9ne10 homo sapien
27	205.5	13.1	217	1 DUSJ_MOUSE	O60966 mus musculu
28	205	13.1	223	1 STYX_MOUSE	O6wuo10 homo sapien
29	204	13.0	223	1 STYX_HUMAN	O02256 saccharomyc
30	202	12.9	364	1 PVH1_YEAST	O9y638 homo sapien
31	201.5	12.9	313	1 M5T1_HUMAN	P51452 homo sapien
32	195.5	12.5	185	1 DUS3_HUMAN	O9d7x3 mus musculu
33	194	12.4	185	1 DUS3_MOUSE	

34	185	11.8	807	1 YB9T_YEAST	P38148 saccharomyc
35	183.5	11.7	276	1 PTP3_CHLEU	O39491 chlamydomon
36	181.5	11.6	278	1 PMP1_SCHPO	O13453 schizosacch
37	178.5	11.4	295	1 DUSF_HUMAN	O9h12 homo sapien
38	177	11.3	125	1 DUSF_MOUSE	O8r42 mus musculu
39	171.5	11.0	339	1 DUSC_MOUSE	O9d012 mus musculu
40	171	10.9	489	1 MSG5_YEAST	P38590 saccharomyc
41	158.5	10.1	340	1 DUSC_HUMAN	O9un16 homo sapien
42	157	10.0	209	1 Y1L3_YEAST	P40479 saccharomyc
43	155	9.9	198	1 DUSD_HUMAN	O9u16 homo sapien
44	152	9.7	292	1 Y042_CAEEL	P34650 caenorhabd
45	150	9.6	198	1 DUSD_MOUSE	O9qy7 mus musculu

ALIGNMENTS

RESULT 1
ID DUSG_HUMAN STANDARD; PRT; 665 AA.
AC Q9BY84; Q9COG3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 16 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 7) (MAP kinase
DE phosphatase 7) (MKP-7).
GN DUSP16 OR MKP7 OR KIAA1700.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
RT "MKP-7, a novel mitogen-activated protein kinase phosphatase,
RT functions as a shuttle protein".
RL J. Biol. Chem. 276:39002-39011(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro".
RL DNA Res. 7:347-355(2000).
CC -1- FUNCTION: Involved in the inactivation of MAP kinases.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC
DR EMBL: AB05156; BAB40814.1; -;
DR EMBL: AB051487; BAB21791.1; ALT_INIT.
DR HSPSP; Q16828; MKP.
DR Genew; HGNC:17909; DUSP16.
DR MIM; 607175; -;
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0005634; C:nucleus; TAS.

```
DR GO:0004721; F:protein phosphatase activity; TAS.
DR GO:0016311; P:dephosphorylation; TAS.
DR GO:0000188; P:inactivation of MAPK; TAS.
DR GO:0004209; P:pleiotyoin B-sensitive MAPK phosphatase nuc. .; TAS.
DR GO:0004204; P:MAPK nucleus export; TAS.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS: PR01764; MARKPHPTASE.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANESE_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR HydroLase; Nuclear protein.
FT DOMAIN 22 137 RHODANESE.
FT DOMAIN 227 289 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 244 244 PHOSPHOCYSTEINE INTERMEDIATE (BY
SQ SEQUENCE 665 AA; 73101 MW; 1BD853FF08460DF CRC64;
SIMILARITY)
```

```
Query Match 99.2%; Score 1552; DB 1; Length 665;
Best Local Similarity 99.3%; Pred. No. 2.5e-125;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 MAHEMIGTQIV-TERLVALLSSTGTEKVLIDSRPFVENVNTHLEANNINCSKLMKRRLO 60
Db 1 MAHEMIGTQIV-TERLVALLSSTGTEKVLIDSRPFVENVNTHLEANNINCSKLMKRRLO 60
QY 61 DKVLITLTHSAKHVYDIDCSQKVVVYDQSSQDVASLSSDCEFLTVLLGKLEKSFNSVHL 120
Db 61 DKVLITLTHSAKHVYDIDCSQKVVVYDQSSQDVASLSSDCEFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLECGKSTVPTCTISOPCLPVANIGPTRILPLVYIGCQRDVINKELI 180
Db 121 LAGGFAEFSRCFPGLECGKSTVPTCTISOPCLPVANIGPTRILPLVYIGCQRDVINKELI 180
QY 181 QONGIGYVLAASYTCRPFDFIESHFLRPVNDSCFCEKILPMLDSVDFIKKASNGCV 240
Db 181 QONGIGYVLAASYTCRPFDFIESHFLRPVNDSCFCEKILPMLDSVDFIKKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAARFVKEKRPITSPNPNFLGQLLDYEKKIKN 300
Db 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAARFVKEKRPITSPNPNFLGQLLDYEKKIKN 300
QY 301 QT 302
Db 301 QT 302
```

RESULT 2

DUS8_HUMAN STANDARD; PRT; 625 AA.

AC 013202;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Dual specificity protein phosphatase hvh-5).
GN DUSP8 OR VH5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96009533; PubMed=7561881;
RA Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.B.;

```
RT "hvh-5: a protein tyrosine phosphatase abundant in brain that
RT inactivates mitogen-activated protein kinase.",
RL J. Neurochem. 65:1823-1833(1995).
CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
CC phosphate as well as with serine/threonine-protein phosphate (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- TISSUE SPECIFICITY: Abundant in brain, heart and skeletal muscle.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR EMBL: U27193; AAA83151.1; -
DR HSSP: Q16829; IMKP.
DR Genew: HGN3:3074; DUSP8.
DR MIM: 602038; -
DR GO:0005737; C:cytoplasm; TAS.
DR GO:0004725; F:protein tyrosine phosphatase activity; TAS.
DR GO:0000188; P:inactivation of MAPK; TAS.
DR GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR PRINTS: PR01764; MARKPHPTASE.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANESE_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW HydroLase; Nuclear protein.

FT DOMAIN 23 138 RHODANESE.
FT DOMAIN 162 430 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 310 550 PRO-RICH.
FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
SQ SEQUENCE 625 AA; 65840 MW; DCBEA1487219666 CRC64;
SIMILARITY)

Query Match 63.2%; Score 988; DB 1; Length 625;
Best Local Similarity 63.1%; Pred. No. 5.2e-77;
Matches 190; Conservative 55; Mismatches 54; Indels 2; Gaps 2;

```
QY 1 MAHEMIGTQIV-TERLVALLSSTGTEKVLIDSRPFVENVNTHLEANNINCSKLMKRRLO 59
Db 1 MAGDRLPRKWDADKASLILRGSGPLVIDSRFVYVNSHTVLSVNICCSKLVKRRLO 60
QY 60 QDKVLITLTHSAKHVYDIDCSQKVVVYDQSSQDVASLSSDCEFLTVLLGKLEKSFNSVH 119
Db 61 QDKVLITLTHSAKHVYDIDCSQKVVVYDQSSQDVASLSSDCEFLTVLLGKLEKSFNSVH 119
QY 120 LLAGGFAEFSRCFPGLECGK-STVPTCTISOPCLPVANIGPTRILPLVYIGCQRDVINKELI 178
Db 121 ILTGFPATFSSCFGLCEGKRALPLMSLQPCLPVPSVGLTILPLVYIGSQDVINKELI 180
QY 179 LIQONGIGYVLAASYTCRPFDFIESHFLRPVNDSCFCEKILPMLDSVDFIKKASNG 238
Db 181 LMTONGIGYVLAASYTCRPFDFIESHFLRPVNDSCFCEKILPMLDSVDFIKKASNG 240
QY 239 CVLVHCLAGISRSATIAIAYIMKRMDSLDEAARFVKEKRPITSPNPNFLGQLLDYEKKIKN 298
Db 239 CVLVHCLAGISRSATIAIAYIMKRMDSLDEAARFVKEKRPITSPNPNFLGQLLDYEKKIKN 298
```

Db 241 QVIVHCLAGISRSATIAIAYIMKTGMSSDDAYRFVDRRSPISPNFGLLEYERTL 300
Qy 299 K 299
Db 301 K 301

RESULT 3
DUS8_MOUSE STANDARD; PRT; 663 AA.
ID DUS8_MOUSE
AC 009112;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Neutonal tyrosine threonine phosphatase 1).
GN DUSP8 OR NTP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=63111565; PubMed=873137;
RA Theodosiou A.M., Rodriguez N.R., Nesbit M.A., Ambrose H.J.,
RA Peterson H., McJellan-Arnold E., Boyd Y., Leverisha M.A., Owen N.,
RA Blake D.J., Ashworth A., Davies K.E.;
RT "A member of the MAP kinase phosphatase gene family in mouse
RT containing a complex trinucleotide repeat in the coding region.";
RL Hum. Mol. Genet. 5:675-684(1996).
CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
CC phosphate as well as with serine/threonine-protein phosphate (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- TISSUE SPECIFICITY: Expressed predominantly in brain and lung.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; X95518; CA64772.1; -.
CC HSSP; Q16828; IMKP.
CC MGD; MGI:106626; Dusp8.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydroxase; Nuclear protein.
FT DOMAIN 23 138 RHODANES.
FT DOMAIN 162 432 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 452 459 POLY-ARG.
FT DOMAIN 555 558 POLY-SER.
FT DOMAIN 559 576 POLY-GLY.

FT DOMAIN 577 600 POLY-SER.
FT DOMAIN 311 552 PRO-RICH.
FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SEQUENCE 663 AA; 68847 MW; 416F429A12C1FA7C CRC64;
SIMILARITY).
SQ
Query Match 63.0%; Score 985; DB 1; Length 663;
Best Local Similarity 62.5%; Pred. No. 1e-76;
Matches 188; Conservative 55; Mismatches 56; Indels 2; Gaps 2;
Qy 1 MAHEMIGTQIV-TERLVALLSGTEKVLIDSPFVPEYNTSHLEAININCSKLMKRRLO 59
Db 1 MADDRLPRKVMADKLLASLLRGPGGPLVIDSRSPFVNSCHVLSVNICCSLTKVRRLO 60
Qy 60 QDKVLTLLIIGHAKKNDIDCSQKVVYVDDSSQDVASLSSDCEFTVTLGKLEKSPSYH 119
Db 61 QGRVTAELIOPATRSQVADTEPDVVVDQSTRDASVLAADSFLLSLKLGCEPQVA 120
Qy 120 LLAGFAEFSPRCFPGCEGKSTLVPT-CISQPLPVANIGPTRLPLNYLGCORVLANKE 178
Db 121 ILTNGFATSSCPFGCEGKRALPMSMSLQCLPVPVGLTRILPHLYLGSQKDVANKD 180
Qy 179 LIOQNGIGVYVLAASYTCPRKDFIPESHFLRPVNDSEFCXILPWLDSVDFIKAKASNG 238
Db 181 LMTQNGISYVLAANSNCPRKDFICESRFMRIPINDVYCEKLLPWLDSIFIDKAKLSSC 240
Qy 239 CVLVHCLAGISRSATIAIAYIMKTGMSSDDAYRFVDRRSPISPNFGLLEYERTL 298
Db 241 QVIVHCLAGISRSATIAIAYIMKTGMSSDDAYRFVDRRSPISPNFGLLEYERTL 300
Qy 299 K 299
Db 301 K 301

RESULT 4
DUS8_HUMAN STANDARD; PRT; 482 AA.
ID DUS8_HUMAN
AC Q9Y6W6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 5) (MAP kinase
DE phosphatase 5) (MKP-5).
GN DUSP10 OR MKP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99321929; PubMed=10391943;
RA Tanoue T., Moriguchi T., Nishida E.;
RT "Molecular cloning and characterization of a novel dual specificity
RT phosphatase, MKP-5.";
RL J. Biol. Chem. 274:11949-11956(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20065165; PubMed=10597297;
RA Theodosiou A., Smith A., Gillieron C., Arkinsteal S., Ashworth A.;
RT "MKP5, a new member of the MAP kinase phosphatase family, which
RT selectively dephosphorylates stress-activated kinases.";
RL Oncogene 18:6981-6988(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, Lung, and Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Datchenko L., Marzella K., Farmer A.A., Rubin G.M., Hong L.,

Query Match	30.0%;	Score 469,	DB 1;	Length 482;
Best Local Similarity	35.9%;	Pred. No. 1.2e-32;		
Matches 107;	Conservative 65;	Mismatches 96;	Indels 30;	Gaps 7;
27	LLIDSPFENVNTSHLEARNINCS-KLMKRLQODKLTWELIQ-HSARKVDIDCSQK	84		
Db	173 VIIDCPFEMVKNKSHIOGVHINCADKISRRLRQGITVLDLISCREGKDSKRIIFSTFE	232		

RESULT 5

ID	DUSA MOUSE	STANDARD,	PRT;	483 AA.
AC	O9CZ9;			
AD	10-OCT-2003 (Rel. 42, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)			
DE	(Mitogen-activated protein kinase phosphatase 5) (MAP kinase phosphatase 5) (MKP-5).			
GN	DUSP10 OR MKP5.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6;			
RX	MEDLINE=20515587; PubMed=11060451;			
RT	Maeda K., Shima H., Kikuchi K., Watanabe Y., Matsuda Y.;			
RT	"Expression and comparative chromosome mapping of MKP-5 genes			
RT	DUSP10/Dusp10".			
RT	Cytogenet. Cell Genet. 90:71-74(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/60; TISSUE=Embryo;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai U., Shinagawa A., Shibata K., Yoshino M., Itoh Y.,			
RA	Aizawa T., Hara A., Fukunishi Y., Komoto H., Adachi J., Fukuda S.,			
RA	Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,			
RA	Kodota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,			
RA	Kueth P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Schintl L.M., Steubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bash G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein S., Hill D., Fletcher C., Fujita K., Gariboldi M.F.,			
RA	Gustinich M., Lyons P., Marchionni L., Mashima U., Mazzarelli U., Mombereys P.,			
RA	Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker G., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontecki S.,			
RA	Hayashizaki Y.;			
PT	"Functional annotation of a full-length mouse cDNA collection."			
PT	Nature 409:685-690(2001).			
CC	-1- FUNCTION: Involved in the inactivation of MAP kinases. Has a			
CC	specificity for the MAPK1/MAPK2/MAPK3/MAPK4 subfamily (By			
CC	similarity).			
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein			
CC	tyrosine + phosphate.			
CC	-1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +			
CC	phosphate.			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).			
CC	-1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.			

CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL; AB037908; BAB17680.1; -
 CC EMBL; AK011995; BAB27966.1; -
 CC HSSP; Q16828; 1MKP.
 CC MGD; MGI:1927070; Dusp10.
 CC InterPro; IPR000340; DS_phosphatase.
 CC InterPro; IPR008343; MAPK_phosph.
 CC InterPro; IPR001763; Rhodanese-like.
 CC InterPro; IPR000387; TYR_phosphatase.
 CC Pfam; PF00782; DSPC; 1.
 CC Pfam; PF00581; Rhodanese; 1.
 CC PRINTS; PR01764; MAPKPHPTASE.
 CC SMART; SM00195; DSPC; 1.
 CC SMART; SM00450; RHOD; 1.
 CC PROSITE; PS50206; RHODANASE_3; 1.
 CC PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 CC PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 CC PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 CC KMW; KMW000000; Nucleolar protein.
 CC FT DOMAIN; 169 286 RHODANASE.
 CC FT ACT_SITE; 386 456 PROTEIN-TYROSINE PHOSPHATASE.
 CC FT ACT_SITE; 409 409 PHOSPHOCYSTEINE INTERMEDIATE (BY
 CC SIMILARITY).
 CC FT CONFLICT; 48 48 A -> T (IN REF. 2).
 CC FT CONFLICT; 113 113 P -> S (IN REF. 2).
 CC FT CONFLICT; 386 386 R -> Q (IN REF. 2).
 CC SQ SEQUENCE 483 AA; 52530 MW; 355156235584689 CRC64;
 Query Match 29.9%; Score 468; DB 1; Length 483;
 Best Local Similarity 35.6%; Pred. No. 1.5e-32;
 Matches 106; Conservative 67; Mismatches 95; Indels 30; Gaps 7;
 QY 27 LILDSRPFEVYNSHILEAININCS-KLMGRLOQDVLTLLEIQ-HSAGKVDIDQSO 84
 DB 174 VITDCRPHEYNHSHIGAVHNCADKISRRLQGGKIVLDLISCREGDSFRKIFSK 233
 QY 85 VVVYDOSQDVASLSDPCFLTVLLGKLEKSFNSVHLIAGGFAEFSRCFPGICEGK----- 139
 DB 234 ILYYDENTNPSRVTPEQPHLYLESIKRGKEPRIVLKGGLSFKQNHGVLCDNSLDQ 293
 QY 140 -----STLVPTGISQCLP-VANIGFRIILPNLYLGGQRDVYLNKELIQNGIGY 187
 DB 294 CREVGAGASMAASMLPQ--SVPTTPDLENALTPILPFLFLGNEQDAQDDDTWQRLNIGY 351
 QY 188 VLNAS-----YTPKPDPIFESHFLRPVNDSCFKILPMLDLSVPIEKAKSNGCVIV 242
 DB 352 VINVTTHLPYHYEKLGF---NYKRLPATDSKNKQNRFFEEAPETIEEHAHOGCKGLI 407
 QY 243 HCLAGISRSATIAIAYIMKMDMSLDEAVRFVKEKPTTISPNFNFLGQLLDYKIKYN 300
 DB 408 HCGAGVSRSAITVIAYLMKTRMTMDAYFVGVGKRRIISPNLNFQGLLEFFEDLNIN 465
 RESULT 6
 DUSE_HUMAN STANDARD; PRT; 381 AA.
 AC Q16828; O75109; Q9BSH6;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
 DE phosphatase 3) (MKP-3) (Dual specificity protein phosphatase PYST1).

GN DUSP6 OR MKP3 OR PYST1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Forebrain;
 RX MEDLINE=96312959; PubMed=8670865;
 RA Groom L.A., Sneddon A.A., Alessi D.R., Dowd S., Keyse S.M.;
 RT "Differential regulation of the MAP, SAP and RK/p38 kinases by PyST1,
 RT a novel cytosolic dual-specificity phosphatase.";
 RL EMOB J. 15:3621-3632(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Liver;
 RX MEDLINE=99077745; PubMed=9858808;
 RA Furukawa T., Yatsuoka T., Yousef E.M., Abe T., Yokoyama T.,
 RA Fukushima S., Soeda E., Hoshi M., Hayashi Y., Sunamura M., Kobari M.,
 RA Horii A.;
 RT "Genomic analysis of DUSP6, a dual specificity MAP kinase phosphatase,
 RT in pancreatic cancer.";
 RL Cytogenet. Cell Genet. 82:156-159(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT VAL-114.
 RC TISSUE=Colon, Kidney, Skin, and Stomach;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg K.H., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein W.J., Ushed T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughlano N.A., Peters G.J., Abismen R.D., Mulhally S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalak U., Smallos D.E.,
 RA Schenck A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 204-347.
 RX MEDLINE=99140289; PubMed=10048930;
 RA Stewart A.E., Dowd S., Keyse S.M., McDonald N.Q.;
 RT "Crystal structure of the MAPK phosphatase PyST1 catalytic domain and
 RT implications for regulated activation.";
 RL Nat. Struct. Biol. 6:174-181(1999).
 CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
 CC family.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q16828-1; Sequence=displayed;
 CC Name=2; Synonyms=DUSP6-ALT;
 CC IsoId=Q16828-2; Sequence=VSP_005137;
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carrincci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S., Krzywinski M.I., Skalek U., Smallue D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
CC family (By similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AK05062; BAB23786.1; -;
DR EMBL: AK009131; BAB26093.1; -;
DR EMBL: BC003869; AA003869.1; -;
DR HSSP: Q16828; IMKP.
DR MGD: MG11914853; Dusp6.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSpC; 1.
DR PRINTS: PR01764; MAPKPHPTASE.
DR SMART: SM00195; DSpC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PSS0206; RHODANES_3; 1.
DR PROSITE: PSS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE: PSS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PSS0054; TYR_PHOSPHATASE_DUAL; 1.
DR DR Hydrolyase.
KW DOMAIN 30 148 RHODANES.
FT DOMAIN 206 381 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 293 293 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
FT CONFLICT 22 22 W -> G (IN REF. 1; BAB26093).
FT CONFLICT 34 34 L -> F (IN REF. 1; BAB26093).
SQ SEQUENCE 381 AA; 42407 MW; 7EALPB154FAD2DA CRC64;

Query Match 29.6%; Score 462.5; DB 1; Length 381;
Best Local Similarity 33.7%; Pred. No. 3.2e-32;
Matches 109; Conservative 64; Mismatches 105; Indels 45; Gaps 7;

Qy 19 LESGTEKVLITDRPFVEYNTSHLEININCSKLMKRLQODKVLITELIQ-SAKHY 77
Db 27 LEIGNEHLLMDCRPGQLYESSHIESINVAIEIMLRQKGNLPRALFTCEPDRDF 86
Qy 78 DIDC-SQKVVVYQSSOD-VASLSSDCLTVLLGKLEKSFNSVHLLAGFAEFSRCPGL 135
Db 87 TRCGTGVTVLVYDENSMDNENGTGSEVLGLLKLKDEBCRAFLVLEGGSFKQAFALH 146

Qy 136 CEKSTLVPTC-ISOPCLPVANIG----- 158
Db 147 CE--TWLDSGSSSPPLVGLGGRISDSSSDIEDLRDPNGATBDSGSPLSNQ 204
Qy 159 --PTRIPLNYLCQDQVINKELIQNGIGYVLAASYCPKP-DPIESHFLRPVNSF 215
Db 205 SFVEILPFLYIGCAQDSTLVDLIEFGIKYILNTVPNLPLENAGDFYKQIPISDH 264
Qy 216 CEKILWLDKSDVFIKARASNCVTVHCLAGISBATAIAYIMKMDSLDEARFVK 275
Db 265 SQNLISQFFPEALSFIDEARKNGVHCLAGISBATAIAYIMKMDSLDEARFVK 324
Qy 276 EKRPISPNPFLGOLLDYEKKI 298
Db 325 MKSNISPNPFLGOLLDYEKKI 347
RESULT 8
DUS6_RAT STANDARD; PRT; 381 AA.
AC 064346;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 6 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 3) (MAP kinase
DE phosphatase 3) (MKP-3).
GN DUSP6 OR MKP3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Neuron;
RX MEDLINE=96216487; PubMed=8631996;
RA Moutrey R.J., Vega Q.C., Campbell J.S., Wenderoth M.P., Hauschka S.D.,
RA Krebs E.G., Dixon J.E.;
RT "A novel cytoplasmic dual specificity protein tyrosine phosphatase
RT implicated in muscle and neuronal differentiation.";
RL J. Biol. Chem. 271:3795-3802(1996).
CC -1- FUNCTION: Inactivates MAP kinases. Has a specificity for the ERK
CC family. Implicated in muscle and neuronal differentiation.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X94185; CAA63895.1; -;
DR EMBL: U42627; AAB06202.1; -;
DR HSSP: Q16828; IMKP.
DR InterPro: IPR000340; DS_phosphatase.

```

QY 19 LBSGTAKVLLIDRPVEYNTHTIILNAININCKMLKRRLLQODKVLITELIQH-SAKHKV 77
Db 27 LELNGEQLLIMCROPELYESSHTIESAINVALPGIMLRLLQXNIPVVALFTRCEDNRIF 86
QY 78 DIDD--SQKVVVYDQSSQD--VASLSSDCEFLVLLIGLTERKSPNSVHLLAGGFMEFRCPPGL 135
Db 87 TTRCGTDVTLVDENSSDMNNTENGESVGLIKKLKDEGCRAPFLBEGFSKFOAEFALH 146
QY 136 CEKGSITVPTC--ISQPCLPVANIG----- 158
Db 147 CE--TNLDGSCSSSPFLPVGLGGLRLISSDSSDIESLDRDNPNSATDSDGSPLSNSQP 204
QY 159 --PRLINLVYIGCRDVLANKELIQONGIGVYLNASTCPRP--DFIESHPFLVPVNDST 215
Db 205 SPPELLLEFLVLGCAKOSTNLDVLEBEFGIKYILVNTENLNLPENNAGEFFKQIPIPSIDHW 264
QY 216 CEKILPWLDCSVDPTEKAKANGCVLVHCLAGISRSKATIAIYIMKRMDSIDEAARYFK 275
Db 265 SQNTSGFEPALISITIDBAKNGCVLVHCLAGISRSVTVVAIYIMOKLINSNDVADIYK 324
QY 276 EKRTTISPNRNFGLQILLDYERKTI 298
Db 325 MKKSNISPNENMGQLLDPERFTL 347

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

[illegible]

ID	DUS4 CHICK	STANDARD	PRT	375 AA.
RESULT 10				
ID	DUS4 CHICK	STANDARD	PRT	375 AA.
AC	Q9PM71			
DT	16-OCT-2001	(Rel. 40. Created)		
DT	16-OCT-2001	(Rel. 40. Last sequence update)		
DT	10-OCT-2003	(Rel. 42. Last annotation update)		
DE	Dual specifically protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)			
DE	(Mitogen-activated protein kinase phosphatase-2) (MAP kinase			
DE	phosphatase-2) (MKP-2).			
GN	DUSP4 OR MKP2.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]_			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=White leghorn;			
RX	MEDLINE=20379393; PubMed=10918612;			
RA	Fu S.-L., Wana A., Vogt P.K.;			
RT	"Identification and characterization of genes upregulated in cells			
RL	transformed by V-Jun.";			
RL	Oncogene 19:3537-3545(2000).			
CC	-I- FUNCTION: Regulates mitogenic signal transduction by			
CC	dephosphorylating both Thr and Tyr residues on MAP kinases ERK1			
CC	and ERK2 (By similarity).			
CC	-I- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein			
CC	tyrosine + phosphate.			

```

CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF167296; AAD46656.1; -.
CC HSSP: Q16828; 1MKP.
CC Interpro: IPR000340; DS phosphatase.
CC Interpro: IPR001763; MAPK phosph.
CC Interpro: IPR001763; Rhodanese-like.
CC Interpro: IPR000387; TYR phosphatase.
CC Pfam: PF00782; DSPC; 1.
CC Pfam: PF00581; Rhodanese; 1.
CC PRINTS: PR01764; MAPKPHPTASE.
CC SMART: SM00195; DSPC; 1.
CC SMART: SM00450; RHOD; 1.
CC PROSITE: PS50206; RHODANESE_3; 1.
CC PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
CC PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
CC PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
CC Hydrobase: Nuclear protein.
CC DOMAIN 25 143 RHODANESE.
CC PT DOMAIN 178 375 PROTEIN-TYROSINE PHOSPHATASE.
CC FT ACT_SITE 261 261 PHOSPHOCYSTEINE INTERMEDIATE (BY
CC SIMILARITY).
CC SO SEQUENCE 375 AA; 41052 MW; 179290D0C2BEEF1 CRC64;
Query Match 27.6%; Score 431; DB 1; Length 375;
Best Local Similarity 33.2%; Pred. No. 1.6e-29;
Matches 104; Conservative 65; Mismatches 120; Indels 24; Gaps 9;
QY 4 EMIGTOIVTERVALLESSTGTEKVLIIISRPFEVNTSHILEATININCSKIMKRRLOQDKV 63
DB 9 EMGGSAL--RLVGRERASGRCILDCRFPLASAHITGALNVRNTTVRR--AKGAV 65
QY 64 LITELI--QHSARKKVDIDSOXVVYDQSSODVASLSDCFLVLLGKLEKSF--NSVH 119
DB 66 SLRGQILPAEGEVARLRLAAGLYTAVLVLDERSPRALRDSVTVALVRLRRDWARADIR 125
QY 120 LLAGGFASFESRCFPGLCGKSTLVPTCISDP-----C-LPVANI-GPTRLLEN 165
DB 126 LLAGGYRFASFEYEPFCAKTKTL--SSISPPSSAESLIDLFSSSCGPTLHDQGGVELLPF 183
QY 166 LVYGCQDVANKELIQNGIGVYNASVYTCPKPDPIPESHPLRPVAVDSFCCKLPLMDK 225
DB 184 LVYGSAYHAARRDMLDALGITLALNVSSDCGN-HFEGHYQYKCIPEVDNKAADISSNFM 242
QY 226 SVDPIEKAKANSQCVLVHCLAGISRSATIAIAYIMKMDSLDEAYRFVKEKPTSPNF 285
DB 243 AIEYIDSVKCCGCVLVHQCAGISRSATICLAYIMMKKRYLLEKAFEFVQGRSIIISPNF 302
QY 286 NPLGQLIDYEKKI 298
DB 303 SFGQLQFESQV 315

```

```

DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
DE 3CH134) (Protein-tyrosine phosphatase ERP).
DE DUSP1 OR PTPN10 OR MKP1 OR 3CH134 OR PTPN16.
DE Mus musculus (Mouse).
DE OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
DE NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BALB/c;
RX MEDLINE=92158357; PubMed=1741163;
RA Charles C.H., Abler A.S., Lau L.F.;
RT "cDNA sequence of a growth factor-inducible immediate early gene and
RT characterization of its encoded protein."
RL Oncogene 7:187-190(1992).
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RX MEDLINE=93360956; PubMed=8355678;
RA Noguchi T., Metz R., Chen L., Mattei M.-G., Carrasco D., Bravo R.;
RT "Structure, mapping, and expression of erp, a growth factor-inducible
RT gene encoding a nontransmembrane protein tyrosine phosphatase, and
RT effect of ERP on cell growth."
RL Mol. Cell. Biol. 13:5195-5205(1993).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=Czech II; TISSUE=Mammary gland;
RX MEDLINE=92388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Heien F.,
RA Diatchenko L., Matrusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedon T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.V., Hulyk S.W.,
RA Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Botterfield Y.S.N., Krzywinski M.I., Skalska U., Smilins D.E.,
RA Schnerch A., Schen J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RN CHARACTERIZATION.
RP MEDLINE=94037096; PubMed=8221888;
RX Sun H., Charles C.H., Lau L.F., Tonks N.K.;
RT "MKP-1 (3CH134), an immediate early gene product, is a dual
RT specificity phosphatase that dephosphorylates MAP kinase in vivo."
RL Cell 75:487-493(1993).
RN [5]
RN FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
RN kinase ERK2 on both Thr-183 and Tyr-185.
RN CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
RN tyrosine + phosphate.
RN -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
RN phosphate.
RN -1- INDUCTION: By growth factors.
RN -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
RN -1- Non-receptor class dual specificity subfamily.
RN -1- SIMILARITY: Contains 1 rhodanese domain.
RN -----
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation-
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.isb-sib.ch/announce/

```


or send an email to license@isb-sib.ch.

CC EMBL; X61940; CAA43944.1; -
 CC EMBL; S64851; AAB27882.1; -
 DR EMBL; BC006967; AAH06967.1; -
 DR PIR; A54681; S24411.
 DR HSSP; Q16828; IMKP.
 DR MGI; MGI:105120; Dusp1.
 DR InterPro; IPR000340; DS phosphatase.
 DR InterPro; IPR008343; MAPK phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; Tyr_phosphatase.
 DR Pfam; PF00782; Dsrc; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; F001764; MAPKPHPTASE.
 DR SMART; SM00195; Dsrc; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PSS0206; RHODANES_3; 1.
 DR PROSITE; PSS0383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PSS0056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolyase; Cell cycle.
 FT DOMAIN 20 137 RHODANES.
 FT ACT_SITE 175 367 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE.
 FT MUTAGEN 258 C->S; LOSS OF ACTIVITY.
 SQ SEQUENCE 367 AA; 39369 MW; 50B5F90FEBBD19AB CRC64;

Query Match 27.2%; Score 425; DB 1; Length 367;
 Best Local Similarity 34.3%; Pred. No. 5e-29;
 Matches 104; Conservative 48; Mismatches 129; Indels 22; Gaps 5;

15 LVALLSSTKLVLDIDSRPFVNTSHILANINNSKMKRLQODKVLTELIOHSAK 74
 13 LRLALREBAACCLLDGSPFAFAGHIAQVAVRSTYVRRRAKAGMGLHIVPAELR 72
 75 HKVDIDGQKVVYVDGSSQDVASISDPCFLVLLGLK--EKPNSVHLLAGFAFSPKCF 132
 73 GRLLAGVYHVAVLIDERSASILDGAKRDGTLAAGLCEHASTQVFTLQGGYEAFASSAC 132
 133 PGLCEGKSTLVPTCTISOP-----CLFVANI-----GPRILLENLYLGGCORDVL 175
 133 PELCSKST--PGLSLPSTLSTVPDASGSCSSCPPLVQDGPPEILSLFYLSGAVHAS 190
 176 NKELIQNGIGYVLNASTYCPKDPFLPESHPLRVPNDSFCEKILPMLDKSVPTERAKA 235
 191 RKMLDALGITLALINVSANCPN-HFEHGYQKSIYEDNHKADISSWFEKADIDPISIKD 249
 236 SNGCVLVHCLAGISRSATTAIYIMKMDMSLDEAYRTVKEKRPITISFNFLGQLIDYE 295
 250 AGGRVFWHQAQGSRSATTCIAVIMKRVKVLDEAFEFVKORRSILISFNFSFGQLQFE 309
 296 KKI 298
 310 SQV 312

RESULT 12
 DUSA4 HUMAN
 ID DUSA4 HUMAN STANDARD; PRT; 394 AA.
 AC 01315; 013524;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
 phosphatase-2) (MKP-2) (Dual specificity protein phosphatase hvm2).
 GN DUSP4 OR MKP2 OR VHZ.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OK NCBI_TaxId=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=95221370; PubMed=7535768;
 RA Guan K.-L., Butch E.;
 RT "Isolation and characterization of a novel dual specific phosphatase,
 RT HVM2, which selectively dephosphorylates the mitogen-activated
 RT protein kinase.";
 RL J. Biol. Chem. 270:7197-7203 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96198119; PubMed=8626452;
 RA Chu Y., Soisak P.A., Khosravi-Far R., Der C.J., Kelly K.;
 RT "The mitogen-activated protein kinase phosphatases PAC1, MKP-1, and
 RT MKP-2 have unique substrate specificities and reduced activity in vivo
 RT toward the ERK2 sevenmaker mutation.";
 RL J. Biol. Chem. 271:6497-6501 (1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin, and Uterus;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Dietchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Abramson J., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Carrinci P., Frange C.,
 RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerker A., Schein J.E., Jones S.J.M., Maiz M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- FUNCTION: Regulates mitogenic signal transduction by
 CC dephosphorylating both Thr and Tyr residues on MAP kinases
 CC ERK1 and ERK2.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U21108; AAA65119.1; -
 DR EMBL; U48807; AAC0452.1; -
 DR EMBL; BC002671; AAH02671.1; -
 DR EMBL; BC014565; AAH14565.1; -
 DR HSSP; Q16828; IMKP.
 DR Genew; HGNC:3070; DUSP4.
 DR MIM; 602747; -
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0008330; P:protein tyrosine/threonine phosphatase acti. .; TAS.
 DR GO; GO:000165; P:MAPKK cascade; TAS.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
 DR InterPro; IPR000340; DS phosphatase.
 DR InterPro; IPR008343; MAPK phosph.
 DR InterPro; IPR001763; Rhodanese-like.


```
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPc; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPc; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANASE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW HydroLase; Nuclear protein.
FT DOMAIN 41 159 RHODANASE.
FT DOMAIN 197 394 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 280 280 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
FT CONFLICT 111 111 R -> G (IN REF. 2).
SQ SEQUENCE 394 AA; 42953 MW; 0603971759B952E CRC64;

Query Match 27.2%; Score 425; DB 1; Length 394;
Best Local Similarity 33.4%; Pred. No. 5,5e-29;
Matches 100; Conservative 64; Mismatches 113; Indels 22; Gaps 9;

QY 19 LBSGTEKVLIDSRPVEYNTSHLEAININCKMKRRLQODKVLITELI--QHSARK 76
DB 39 LPSG-GKCLLDPRFLASAGYILGSVNVRCNTIYRR-AKGSVSLQILPAEEVRR 96
QY 77 VDIDCGQKVVVYDSSQDVASLSDCLFVTLGKLEKSF--NSVHLAGFAFRCFPG 134
DB 97 LRSGLYSANIVYDERSPRASLREDSTVSLVQALRNARNTDILCKGGERFSSEYPE 156
QY 135 LCEGKSTL-----VPTCISQ-----C-LPVANI-GPTRIILPNLYLGCQRDVINKEL 179
DB 157 FCGKTKALAIIPRPVPSPATEPDLGSSCGTPLHDGCGVELLPYLIGSAHAARRDM 216
QY 180 IQQNGIGVYVNASYTCRKPDPFIPESHFLRVVNDSECEKILPWLDSVDFIEKAKSNGC 239
DB 217 LDALGITALLNVSSDCPN-HFECHGYQKCIPEVDNHNKADISSWFEAIEYIDAVKCRGR 275
QY 240 VLVHCLAGISRSATIAIYIMKMDMSLDEARYRVEKRTTISPNRFLGQLLDYEKKI 298
DB 276 VLVHCAQGISRSATICLAYIMMKRVLEAFEFVKORRSIISPNFSFMQLQFESQV 334

RESULT 13
DUS4_RAT STANDARD; PRT; 395 AA.
ID DUS4_RAT 062767;
AC 062767;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase-2) (Map kinase
DE phosphatase-2) (MKP-2).
GN DUSP4 OR MKP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Phenochromocytoma;
RX MEDLINE=95301550; Pubmed=7782322;
RA Miera-Press A., Rim C.S., Yao H., Robertson M.S., Stork P.J.S.;
RT "A novel mitogen-activated protein kinase phosphatase. Structure,
RT expression, and regulation."
RL J. Biol. Chem. 270:14587-14596(1995).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases ERK1
CC and ERK2 (by similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
```

```
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC TISSUE SPECIFICITY: Expressed at moderate levels in nearly all
CC tissues and cells including brain, spleen, and testes with the
CC higher expression in the heart and lung and lower expression in
CC skeletal muscle and kidney. Undetectable in liver. Expressed in
CC many areas of the brain with very strong expression in the
CC hippocampus, piriform cortex, and the suprachiasmatic nucleus.
CC -1- INDUCTION: By mitogens and by stress.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23438; AAC52493.1; -.
DR HSSP; Q16828; IMKP.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPc; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPc; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANASE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW HydroLase; Nuclear protein.
FT DOMAIN 42 160 RHODANASE.
FT DOMAIN 198 395 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 281 281 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 395 AA; 43187 MW; A90EFPD378A050FD CRC64;

Query Match 27.1%; Score 423.5; DB 1; Length 395;
Best Local Similarity 32.2%; Pred. No. 7.4e-29;
Matches 98; Conservative 61; Mismatches 118; Indels 27; Gaps 7;

QY 16 VALLBSGTEKVLIDSRPVEYNTSHLEAININCKMKRRLQODKVLITELI--QHSAR 73
DB 38 LGLLSGG--KCLLDPRFLASAGYILGSVNVRCNTIYRR-AKGSVSLQILPAEEV 94
QY 74 KHKVDIDCGQKVVVYDSSQDVASLSDCLFVTLGKLEKSF--NSVHLAGFAFRCFPG 131
DB 95 RARLRGLYSANIVYDERSPRASLREDSTVSLVQALRNARNTDILCKGGERFSSE 154
QY 132 FPGELGKSTLVPITCISQCLPVANI-----GPTRIILPNLYLGCQRDV 174
DB 155 YEFSCSKTAL--AIPRPVPSPATNESLDGSSCGTPLHDGCGPEILPFLIGSAYRA 212
QY 175 LNKELIQQNGIGVYVNASYTCRKPDPFIPESHFLRVVNDSECEKILPWLDSVDFIEKAK 234
DB 213 AARDMDALGITALLNVSSDCPN-HFECHGYQKCIPEVDNHNKADISSWFEAIEYIDAVK 271
QY 235 ASNGCYLVHCLAGISRSATIAIYIMKMDMSLDEARYRVEKRTTISPNRFLGQLLDY 294
DB 272 DCRGRVLVHCAQGISRSATICLAYIMMKRVLEAFEFVKORRSIISPNFSFMQLQF 331
QY 295 EKKI 298
DB 332 ESQV 335

RESULT 14
DUS2_HUMAN
```

```

ID DUS2_HUMAN STANDARD; PRT; 314 AA.
AC Q05923;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 2 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Dual specificity protein phosphatase PAC-1).
GN DUSP2 OR PAC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93206122; PubMed=7681221;
RA Rohan P., Davis P., Moskaluk C.A., Kearns M., Krutzsch H.,
RT Siebenlist U., Kelly K.,
RL "PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase,"
RN Science 259:1763-1766 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96070437; PubMed=7590752;
RA Yi H., Morton C.C., Weremowicz S., McBride O.W., Kelly K.,
RT "Genomic organization and chromosomal localization of the DUSP2 gene,
RN encoding a MAP kinase phosphatase, to human 2p11.2-q11."
RN Genomics 28:92-96 (1995).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=B-cell;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RT MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schaller G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Blatt N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sepleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fehey J., Helton B., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RN human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Regulates mitogenic signal transduction by
CC dephosphorylating both Thr and Tyr residues on MAP kinases
CC ERK1 and ERK2.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: In hematopoietic tissues.
CC -1- INDUCTION: By mitogens.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; L11329; AAA50779.1; -.

```

```

DR EMBL; U23853; AAA6112.1; -.
DR EMBL; BC007771; AAA6771.1; -.
DR PIR; A57126; A57126.
DR PDB; 1IKZ; 30-MAY-02.
DR Genem; HGNC:3068; DUSP2.
DR MIM; 603068; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; TAS.
DR GO; GO:0008330; F:protein tyrosine/threonine phosphatase acti. .; TAS.
DR GO; GO:0000188; P:inactivation of MAPK; TAS.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR InterPro; IPR003440; DS_phosphatase.
DR InterPro; IPR003443; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHRASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANESE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase; Nuclear protein; 3D-structure.
FT DOMAIN 23 144 RHODANESE.
FT DOMAIN 237 302 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 257 257 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 314 AA; 34399 MW; F0D3543C6DE10C45 CRC64;
Query Match 26.7%; Score 418; DB 1; Length 314;
Best Local Similarity 36.0%; Pred. No. 1.6e-28;
Matches 109; Conservative 45; Mismatches 105; Indels 44; Gaps 8;
QY 24 EKVLLDPSPPVEYNTSHILEININCSKLMKR-----IQDDKLTILELQHS 72
DB 25 ERTLLDCRPLFAFCRRHVARAPVPMALRRRARGPAAVLACLPDRRLRLTVRG 84
QY 73 AAKKVDIDSCQVYVYDSSQDVASLSD---CELVTLGKLEKSPSVYLLAGPAEF 128
DB 85 L-----AAVVLIDGSGASVVELRPDSPAHYLALHETRAGP7AVYFLRGSGDGF 135
QY 129 SRCPGPGLC-EGKSTLVP-----TCISQPCLPVANI-GPTRLTPLYGCGQDVLNKLQ 161
DB 136 QCCCECDLSEAPAPALPPTGKTRSDSRAPVYDGGVELLPYLFGLGSCSHSSDLQGLQ 195
QY 182 QNGIGVYVNASTYCPKDPFIPESHF-----LRVPVDSFCEKILPWLDRKVDPIEKAKA 235
DB 196 ACGITAAVNVASACP-----NHFEGLFRYKSIPEVDNQWEISAMFGQALGFIWMYKX 248
QY 236 SNGCVLVHCLAGISRSATIAIYIMKRMDSLDRAKPYFVEXKPTTSPNPNFGLDLYE 295
DB 249 SSGRVLVHCQAGISRSATICLAIYMQSRVRLDEAFVQKRGVISPNSFNGQLLOFE 308
QY 296 KKI 298
DB 309 TQV 311
RESULT 15
DUS1_HUMAN STANDARD; PRT; 367 AA.
AC P28562;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP kinase phosphatase-1) (MKP-1) (Protein-tyrosine phosphatase
DE C1100) (Dual specificity protein phosphatase hvh1).
DE DUSP1 OR PTPN10 OR MKP1 OR C1100 OR VH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Forebrain;
 RX MEDLINE=93024952; PubMed=1406996;
 RA Keyes S.M., Emalie E.A.;
 RT "Oxidative stress and heat shock induce a human gene encoding a
 RT protein-tyrosine phosphatase.";
 RL Nature 359:644-647(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
 RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
 RA Fahy J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Botterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Dual specificity phosphatase that dephosphorylates MAP
 CC kinase ERK2 on both Thr-183 and Tyr-185.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- INDUCTION: By oxidative stress and heat shock.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL: X68277; CAA48338.1; -.
 CC EMBL: BC022463; AAH2463.1; -.
 DR PIR: S29090; S29090.
 DR HSP: Q16828; IMKP.
 DR GeneW: HGNC:3064; DUSP1.
 DR MIM: 600714; -.
 DR GO: GO:0004726; P:non-membrane spanning protein tyrosine phos. .; TAS.
 DR GO: GO:0006979; P:response to oxidative stress; TAS.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00195; DSPC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PSS0206; RHODANES_3; 1.
 DR PROSITE: PSS0383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PSS0056; TYR_PHOSPHATASE_2; 1.

DR PROSITE: PSS0054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolase; Cell cycle-
 FT DOMAIN 20 137 RHODANES.
 FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 258 258 PHOSPHOCYSTEINE INTERMEDIATE (BY
 FT SIMILARITY).
 SO SEQUENCE 367 AA; 39297 MW; 11BD1D39A9FCD51F CRC64;
 Query Match 26.5%; Score 415; DB 1; Length 367;
 Best Local Similarity 33.9%; Pred. No. 3.6e-28;
 Matches 102; Conservative 47; Mismatches 134; Indels 18; Gaps 4;
 QY 15 LVALLSGTEKYLIDSRPFVEVNTSHILEAININCSKLMKRLQODKYLITELIHSK 74
 DB 13 LRALDGERAAQCLLDRCSEFFAFNAGHIGSVNVRSTIVRRRAKAMGLEHTVPAE 72
 QY 75 HKVDIDCSQKVVVYDQSSQDVASLSDCFLVLLGRL--EKSPNVHLLAGGFAEFSRC 132
 DB 73 GRLLAGAYHAVVLLDERSAALDQAKRDGTLLAAGLCREAPAAQVFLKGGYEAFSASC 132
 QY 133 PGLCEK-----STLVPTCSQPLPVANT-----GPRILPNLYLGGGRVLR 177
 DB 133 PELCSKQSTPMGLSLPLSTVSDSAESSGSCSTPLDQGGPEILPFLYLSAYHASR 192
 QY 178 ELIQNGIGVNLNASTYCPKPDFIPESHFLRVVNDSPCEKILPWLDKSVDPIEKAKSN 237
 DB 193 DMIDALGITRLINVSANCN-HFEGHYQYSIPVEDNHKADISSWNEAIDPDISKNG 251
 QY 238 GCVLVHCLAGISRSATIALAYIMKRMDSLDEAYRPFKEKRPITISFNFLQGLDYEK 297
 DB 252 GRVFVHCQAGISRSATICLAYIMRTRVKLDFAEFVKKRRSISNFSFMGQLQFESQ 311
 QY 298 I 298
 DB 312 V 312

Search completed: June 21, 2004, 13:21:07
 Job time : 6.55843 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:15:22 ; Search time 23.7353 Seconds

(without alignments)
4014.554 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMIGTIVTERLVALL.....PNNFPLGQLDYKKIKNQF 302

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_plant:*
10: sp_rodent:*
11: sp_virus:*
12: sp_vertebrate:*
13: sp_unclassified:*
14: sp_virus:*
15: sp_bacterioplasmid:*
16: sp_bacteriophage:*
17: sp_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1552	99.2	662	4	Q81VT8
2	1545	98.8	665	4	Q6N4Y9
3	1517	97.0	660	11	Q920R2
4	1514	96.8	677	11	Q99MG5
5	1375.5	87.9	355	11	Q9AG16
6	1312	83.9	616	4	Q8N5T1
7	1172.5	75.0	622	11	Q99MG5
8	1034	66.1	300	11	Q9AG15
9	989	63.2	625	4	Q86S88
10	985	63.0	665	11	Q7TS29
11	891	57.0	206	11	Q8BZE4
12	610	39.0	143	4	Q96Q82
13	563.5	36.0	657	5	Q8ST19
14	472.5	30.2	367	13	Q7SZF3
15	469	30.0	483	11	Q8R3L3
16	465	29.7	382	13	Q7T2L9

17	450	28.8	606	5	Q8ST18	Q8ST18 caenorhabdi
18	447	28.6	388	13	Q8UM48	Q8UM48 fuqu rubrip
19	440.5	28.2	368	4	Q8NPU0	Q8NPU0 homo sapien
20	436	27.9	382	13	Q7T2L8	Q7T2L8 brachydantio
21	430	27.5	362	13	Q803B2	Q803B2 brachydantio
22	425	27.2	411	4	Q13649	Q13649 homo sapien
23	423.5	27.1	378	13	Q91663	Q91663 xenopus lae
24	423.5	27.1	398	11	Q8BFW3	Q8BFW3 mus musculu
25	421.5	27.0	369	13	Q91790	Q91790 xenopus lae
26	419.5	26.8	369	13	Q90W58	Q90W58 xenopus lae
27	416	26.6	367	11	Q64193	Q64193 rattus sp.
28	413	26.4	318	11	Q80ZM1	Q80ZM1 mus musculu
29	413	26.4	367	11	Q63683	Q63683 rattus norv
30	411	26.3	439	5	Q81G35	Q81G35 caenorhabdi
31	403.5	25.8	353	13	Q42253	Q42253 gallus gall
32	403.5	25.8	360	13	Q7ZVL8	Q7ZVL8 brachydantio
33	403	25.8	436	11	Q99KC2	Q99KC2 mus musculu
34	397	25.4	452	11	Q8K1S9	Q8K1S9 mus musculu
35	397	25.4	452	11	Q7TNL7	Q7TNL7 mus musculu
36	382.5	24.5	354	4	Q8N4A4	Q8N4A4 homo sapien
37	347.5	22.2	411	5	Q95SV1	Q95SV1 drosophila
38	347.5	22.2	411	5	Q9VVM5	Q9VVM5 drosophila
39	319	20.4	303	4	Q9NSW1	Q9NSW1 homo sapien
40	306	19.6	177	11	Q9CSL5	Q9CSL5 mus musculu
41	287	18.4	369	5	O44128	O44128 caenorhabdi
42	285.5	18.3	476	5	O46122	O46122 drosophila
43	280.5	17.9	476	5	O9VHV8	O9VHV8 drosophila
44	272	17.4	290	5	O86P14	O86P14 drosophila
45	269	17.2	241	5	O9VVM4	O9VVM4 drosophila

ALIGNMENTS

RESULT 1
ID Q81VT8 PRELIMINARY; PRT; 662 AA.
AC Q81VT8;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
DE Similar to dual specificity phosphatase 16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ductennum;
RA Strauberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC042101; AAH42101.1; -
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:Protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE_3; 1.
DR PROSITE; PS00363; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 662 AA; 72818 MW; 350534EF0652B98F CRC64;
Query Match 99.2%; Score 1552; DB 4; Length 662;
Best Local Similarity 99.3%; Pred. No. 4.6e-138;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MAHEMIGTQIVTERLVALLESCTEKVLLIDSRPFVEYNTSHILEAININCSKIMKRRLQ 60
DB 1 MAHEMIGTQIVTERLVALLESCTEKVLLIDSRPFVEYNTSHILEAININCSKIMKRRLQ 60
QY 61 DKVLTITELIQHSANRKHVDIDCSQKVVYVYDSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLTITELIQHSANRKHVDIDCSQKVVYVYDSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCGKSTLVPTCISOPCLPVANIGPTRILPMLYLGCGQDVLNKELI 180
DB 121 LAGGFAEFSRCFPGLCGKSTLVPTCISOPCLPVANIGPTRILPMLYLGCGQDVLNKELI 180
QY 181 QONGIGYVLANASTCPKDPFIPESHFLRVVNDSCFCEKILPMLDKSVDFTEKAKASNGCV 240
DB 181 QONGIGYVLANASTCPKDPFIPESHFLRVVNDSCFCEKILPMLDKSVDFTEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

```

RESULT 2
096N49 PRELIMINARY; PRT; 665 AA.

```

ID 096N49
AC 096N49,
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ31411.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Taahito H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Okunuki T., Saito K., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Nagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima K., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Maeno Y., Nagai K.,
RA Isegai T.;
RT "MEDO human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055973; BAB71060.1; -.
DR HSSP; Q16828; IMKP.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017017; P:MAP kinase phosphatase activity; IEA.
DR GO; GO:006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein; Hydrolase.
SQ SEQUENCE 665 AA; 73058 MW; 1EABDDPF08460DPF CRC64;

```

Query Match 98.8%; Score 1545; DB 4; Length 665;
Best Local Similarity 99.0%; Pred. No. 2.1e-137;
Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MAHEMIGTQIVTERLVALLESCTEKVLLIDSRPFVEYNTSHILEAININCSKIMKRRLQ 60
DB 1 MAHEMIGTQIVTERLVALLESCTEKVLLIDSRPFVEYNTSHILEAININCSKIMKRRLQ 60
QY 61 DKVLTITELIQHSANRKHVDIDCSQKVVYVYDSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLTITELIQHSANRKHVDIDCSQKVVYVYDSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCGKSTLVPTCISOPCLPVANIGPTRILPMLYLGCGQDVLNKELI 180
DB 121 LAGGFAEFSRCFPGLCGKSTLVPTCISOPCLPVANIGPTRILPMLYLGCGQDVLNKELI 180
QY 181 QONGIGYVLANASTCPKDPFIPESHFLRVVNDSCFCEKILPMLDKSVDFTEKAKASNGCV 240
DB 181 QONGIGYVLANASTCPKDPFIPESHFLRVVNDSCFCEKILPMLDKSVDFTEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITISPNFNLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

```

RESULT 3
0920R2 PRELIMINARY; PRT; 660 AA.

```

ID 0920R2
AC 0920R2,
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE MAP kinase phosphatase-7.
GN DUSP16 OR 3830417M17RK OR MKP-7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Cerebellum;
RX MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shima H., Watanabe M., Kikuchi K.,
RT "MKP-7, a Novel Mitogen-activated Protein Kinase Phosphatase,
RT Functions as a Shuttle Protein."
RL J. Biol. Chem. 276:39002-39011(2001).
DR EMBL; AB052157; BAB47240.1; -.
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0000188; F:inactivation of MAPK; IDA.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; kinase.
SQ SEQUENCE 660 AA; 72695 MW; DB609FCDDAD4AA309 CRC64;

```

Query Match 97.0%; Score 1517; DB 11; Length 660;
Best Local Similarity 96.0%; Pred. No. 9.3e-135;
Matches 290; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 MAHEMIGTOIVTERIVALLESGTEKVLIDSRPVEYNTSHLEAININCSKLMKRRLQ 60
DB 1 MAHEMIGTOIVTERIVALLESGTEKVLIDSRPVEYNTSHLEAININCSKLMKRRLQ 60
QY 61 DKVLITTELOHSAKHKVDIDCSQKVVYDQSSODVASLSSDCEFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLITTELOHSAKHKVDIDCSQKVVYDQSSODVASLSSDCEFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCPGLCEGKSTLVPTCISQCLPVANIGPTRILPNIYLGCGQDVLNKL 180
DB 121 LAGGFAEFSRCPGLCEGKSTLVPTCISQCLPVANIGPTRILPNIYLGCGQDVLNKL 180
QY 181 QONGIGYVLANASTCPKPDPIESHFLRPVNDSPCEKILPWLKSDVDFTEKAKSNGCV 240
DB 181 QONGIGYVLANASTCPKPDPIESHFLRPVNDSPCEKILPWLKSDVDFTEKAKSNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTISPNFNLGQLDYKCKIN 300
DB 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTISPNFNLGQLDYKCKIN 300
QY 301 QT 302
DB 301 QT 302
```

```
RESULT 4
Q99MG6 PRELIMINARY; PRT; 677 AA.
AC Q99MG6;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Map kinase phosphatase-M A1 isoform.
GN DUSP16 OR 3830417M17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RA Yoshikai Y.,
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages."
RL Mol. Cell. Biol. 20:6999-7009(2001).
DR EMBL: AF345951; AAK35052.1; -.
DR HSSP: Q16828; IMKP.
DR MCD; MG1:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MARKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PSS0206; RHODANSE_3; 1.
DR PROSITE; PSS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PSS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
KW HydroLase; Kinase.
SQ SEQUENCE 677 AA; 74550 MW; 8B6D5B7096CEC2FC CRC64;
```

Query Match 96.8%; Score 1514; DB 11; Length 677;
Best Local Similarity 95.7%; Pred. No. 1,9e-134;
Matches 289; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 MAHEMIGTOIVTERIVALLESGTEKVLIDSRPVEYNTSHLEAININCSKLMKRRLQ 60
DB 1 MAHEMIGTOIVTERIVALLESGTEKVLIDSRPVEYNTSHLEAININCSKLMKRRLQ 60
QY 61 DKVLITTELOHSAKHKVDIDCSQKVVYDQSSODVASLSSDCEFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLITTELOHSAKHKVDIDCSQKVVYDQSSODVASLSSDCEFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCPGLCEGKSTLVPTCISQCLPVANIGPTRILPNIYLGCGQDVLNKL 180
DB 121 LAGGFAEFSRCPGLCEGKSTLVPTCISQCLPVANIGPTRILPNIYLGCGQDVLNKL 180
QY 181 QONGIGYVLANASTCPKPDPIESHFLRPVNDSPCEKILPWLKSDVDFTEKAKSNGCV 240
DB 181 QONGIGYVLANASTCPKPDPIESHFLRPVNDSPCEKILPWLKSDVDFTEKAKSNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTISPNFNLGQLDYKCKIN 300
DB 241 LVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTISPNFNLGQLDYKCKIN 300
QY 301 QT 302
DB 301 QT 302
```

```
RESULT 5
Q9AG16 PRELIMINARY; PRT; 355 AA.
AC Q9AG16;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Map kinase phosphatase-M B1 isoform.
GN DUSP16 OR 3830417M17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RA Yoshikai Y.,
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages."
RL Mol. Cell. Biol. 20:6999-7009(2001).
DR EMBL: AF345953; AAK35054.1; -.
DR HSSP: Q16828; IMKP.
DR MCD; MG1:1917936; Dusp16.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS phosphatase.
DR InterPro: IPR008343; MAPK_phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MARKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PSS0206; RHODANSE_3; 1.
DR PROSITE; PSS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PSS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
KW HydroLase; Kinase.
SQ SEQUENCE 355 AA; 39502 MW; D52C29AE215CA285 CRC64;
```

Query Match 87.9%; Score 1375.5; DB 11; Length 355;
Best Local Similarity 94.6%; Pred. No. 1e-121;
Matches 265; Conservative 8; Mismatches 4; Indels 3; Gaps 1;

250 IENAKASNGCVLHVHCLAGISRSATIAIAYIMKRDMSLDEAYRFVKEKPTISPENFLG 289

Query Match	75.0%	Score 1172.5	DB 11	Length 622
Best Local Similarity	77.5%	Pred No. 3.5e-102		
Matches 234	Conservative 8	Mismatches 5	Indels 55	Gaps 1

Y	1	MAHNMGTQVYTRRLVALLSGTEKVVLLIDSRPFVEYNTHIIEAININCSKIMRRLOQ	60
C	1	MAHNMGTQVYESTVALLSGTEKVVLLIDSRPFVEYNTHIIEAININCSKIMRRLOQ	60
Y	61	DYVLTIELIQHSARKVVIDCSQKVVVYDQSDQVAVLSIDCEFLVYLIDKTEKSNVHL	120
C	61	DYVLTIELIQHSARKVVIDCSQKVVVYDQSDQVAVLSIDCEFLVYLIDKTEKSNVHL	120
Y	121	LAGGFAPFSRCPGLCEGKSTLVPTGISQPCLPVANI GPRILIPNLYLGCQDQVINKELI	180
C	121	LA-----DLM	125
Y	181	QONGIGYVTLNASTYCPKDPILPESHFLRVVNDSPCEKILPMWIDKSVDFIEKAKASNGCV	240
C	126	QONGIGYVTLNASTYCPKDPILPESHFLRVVNDSPCEKILPMWIDKSVDFIEKAKASNGCV	185
Y	241	LVHCTAGISRSATTAIAVIMKRMMSIDEAYRPFVKERPTISPNRNFQGLLDVEKKIKN	300

Db 186 LIHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPTISPNFNFQGLMDYEKTINN 245
QY 301 QT 302
Db 246 QT 247

RESULT 8
ID Q9AG15 PRELIMINARY; PRT; 300 AA.
AC Q9AG15;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Map kinase phosphatase-M B2 isoform.
GN DUSP16 OR 3830417M1RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
RA Yoshikai Y.;
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
RT Activation in Macrophages";
RL Mol. Cell. Biol. 20:6999-7009(2001).
DR EMBL; AF345954; AAK35055.1; -.
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1917936; C:cytoplasm; IDA.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro; IPR000340; D5 phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SM00450; RHOD; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydrolyase; Kinase.
SQ SEQUENCE 300 AA; 33624 MW; 4C61846ACDF0F456 CRC64;

Query Match 66.1%; Score 1034; DB 11; Length 300;
Best Local Similarity 75.0%; Pred. No. 1.8e-89;
Matches 210; Conservative 8; Mismatches 4; Indels 58; Gaps 2;

QY 1 MAHEMIGTQVTERLVALLSSTGTFKLLIDSRPFVENVNTSHILEAININCSKLMKRLQ 60
Db 1 MAHEMIGTQVTERLVALLSSTGTFKLLIDSRPFVENVNTSHILEAININCSKLMKRLQ 60
QY 61 DKVLITELIHSAGKAVDIDCSQKVVVYDDOSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
Db 61 DKVLITELIHSAGKAVDIDCSQKVVVYDDOSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIIGPTRILPNLYLGGQRDVLANKEI 180
Db 121 LA-----DLN 125
QY 181 QQNGIGVVLNASYTCPRDPFIPESHFLRPVNDSEFCETILPMLDKSVDFLEKASNGCV 240
Db 126 QQNGIGVVLNASYTCPRDPFIPESHFLRPVNDSEFCETILPMLDKSVDFLEKASNGCV 185
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPT 280
Db 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPT 280

Db 186 LIHCLAGISRSATIAIAYIMKMDMSLDEAYR---RQKPT 222

RESULT 9
ID Q86SS8 PRELIMINARY; PRT; 625 AA.
AC Q86SS8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to dual specificity phosphatase 8.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Astrocytoma;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045110; AA045110.1; -.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; D5 phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR002965; P rich extensn.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR PRINTS; PRO1217; PRICEXTENS.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 625 AA; 65826 MW; C7C608407B724FPC CRC64;

Query Match 63.2%; Score 989; DB 4; Length 625;
Best Local Similarity 63.1%; Pred. No. 8.3e-85;
Matches 190; Conservative 55; Mismatches 54; Indels 2; Gaps 2;

QY 1 MAHEMIGTQV-TERLVALLSSTGTFKLLIDSRPFVENVNTSHILEAININCSKLMKRLQ 59
Db 1 MAGDRIPKRVMDKKLASLRLGGPGPLVYDSSQVFNYSWHLSSVNICCSKLVKRLQ 60
QY 60 QDVLITELIHSAGKAVDIDCSQKVVVYDDOSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 119
Db 61 QGKVTIELIHSAGKAVDIDCSQKVVVYDDOSSQDVASLSSDCFLTVLLGKLEKSFNSVHL 120
QY 120 LLAGGFAEFSRCFPGCEGK-STLVPTCISQPCLPVANIIGPTRILPNLYLGGQRDVLANKE 178
Db 121 ILTGCAETSSCPFGCEGKPALILPMSLSQPCLPVPSVGLTRILPHLYLGSQKDVLANKD 180
QY 179 LLOQNGIGVVLNASYTCPRDPFIPESHFLRPVNDSEFCETILPMLDKSVDFLEKASNG 238
Db 181 LMTQNGISVVLNASYTCPRDPFIPESHFLRPVNDSEFCETILPMLDKSVDFLEKASNG 240
QY 239 CVLVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPTISPNFNFQGLMDYEKTI 298
Db 241 QVTVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPTISPNFNFQGLMDYEKTI 300
QY 299 K 299
Db 301 K 301

RESULT 10
Q7TS29 PRELIMINARY; PRT; 665 AA.
ID Q7TS29;
AC Q7TS29;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TRENBLREL. 25, last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, last annotation update)
 DE Dusp8 protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heist L.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butlerfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Maitra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL EMBL: BC052705; AAHS2705.1; -
 SQ SEQUENCE 665 AA; 69021 MW; 9166E36A8835249F CRC64;

Query Match 63.0%; Score 985; DB 11; Length 665;
 Best Local Similarity 62.5%; Pred. No. 2,2e-84;
 Matches 188; Conservative 55; Mismatches 56; Indels 2; Gaps 2;

QY 1 MAHEMGTQIV-TERLVALLSGTEKVLIDSRPFVEYNTSHILEININSGKMKRRLOQ 59
 DB 1 MADGRLEPRKVMKAKKLASLRGSGPLVYIDSRPFVEYNSCHVLSVAVICSKYKRRLOQ 60

QY 60 QDKVLTTELIOHSAKHKVIDCSOKVYVYDSSQVAVSSDCFLTVLIGLERSFNSVH 119
 DB 61 QGKVTIAELIOPATRSQVATEPQDVVYDOSTRASVLAADSFSLISLKDCCFDSVA 120

QY 120 LLAGFAEFSRCFPGICEGKSTLVPT-CISQPCLPVANIPTRIIPNLVYGGQDVYANK 178
 DB 121 ILTGFAPFPSSCFPGICEGKSTLVPTLPSMGLPVPSCVGLTRILPHLYLSQSDVYANKD 180

QY 179 LTIQNGIGVYLVASTYCKPDPFIPESHRLRPVNDSEFEKTLPLWLDKSVDFIEKAXASNG 238
 DB 181 LMTQNGIGVYLVANASCRKPDPTCESRFRIPINDNYCEKTLPLDKSIFIDIKAKLSsc 240

QY 239 CVLVCLAGISRSATIALAYIMKRMDSLDEAYRFEVEKPTTSPNPNFGOLLDYEKXI 298
 DB 241 QVIVACLAGISRSATIALAYIMKMGWSSDAYFVDRRPSISPNENFLGOLLEYERSL 300

QY 299 K 299
 DB 301 K 301

RESULT 11
 Q8BZE4 PRELIMINARY; PRT; 206 AA.
 AC Q8BZE4;
 DT 01-MAR-2003 (TRENBLREL. 23, Created)

DT 01-MAR-2003 (TRENBLREL. 23, last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, last annotation update)
 DE Truncated MAPK phosphatase 7 homolog.
 GN DEERD213E
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=urinary bladder;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL: AK035652; BAC29138.1; -
 DR MGD: MGI:1098836; DEERD213E.
 DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR GO: GO:0006470; P:Protein amino acid dephosphorylation; IEA.
 DR InterPro: IPR001763; Rhodanese-like.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00450; RHODANESE_1.
 DR PROSITE: PSS0206; RHODANESE_3; 1.
 SQ SEQUENCE 206 AA; 22566 MW; E7PDP5C01ADC2F73 CRC64;

Query Match 57.0%; Score 891; DB 11; Length 206;
 Best Local Similarity 97.7%; Pred. No. 3.8e-76;
 Matches 173; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILEININSGKMKRRLOQ 60
 DB 1 MAHEMGTQIVTESVLALLSGTEKVLIDSRPFVEYNTSHILEININSGKMKRRLOQ 60

QY 61 DKVLITELIOHSAKHKVIDCSOKVYVYDSSQVAVSSDCFLTVLIGLERSFNSVH 120
 DB 61 DKVLITELIOHSAKHKVIDCSOKVYVYDSSQVAVSSDCFLTVLIGLERSFNSVH 120

QY 121 LAGFAEFSRCFPGICEGKSTLVPTCISQPCLPVANIPTRIIPNLVYGGQDVYANK 177
 DB 121 LAGFAEFSRCFPGICEGKSTLVPTCISQPCLPVANIPTRIIPNLVYGGQDVYANK 177

RESULT 12
 Q96Q82 PRELIMINARY; PRT; 143 AA.
 AC Q96Q82;
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, last annotation update)
 DE Truncated MAPK phosphatase 7.
 GN MKP7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Montpelit A., Bolly G., Simmet D.,
 RT "A detailed transcriptional map of the chromosome 12p12 tumor
 RT suppressor locus.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY038927; AAK69770.1; -
 DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR GO: GO:0006470; P:Protein amino acid dephosphorylation; IEA.
 DR InterPro: IPR008143; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00450; RHODANESE_1.
 DR PROSITE: PSS0206; RHODANESE_3; 1.

DR PROSITE: PS50206; RHODANESE_3; 1.
SQ SEQUENCE 143 AA; 16077 MW; 5213A213AA7E5974 CRC64;
Query Match 39.0%; Score 610; DB 4; Length 143;
Best Local Similarity 98.4%; Pred. No. 9.5e-50;
Matches 124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAHEMGTQVTEERLVALLESSTGKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
DB 1 MAHEMGTQVTEERLVALLESSTGKYLIDSRPFVEYNTSHLEAININCSKLMKRLQ 60
QY 61 DKYLITELIQSHAKHKVIDCSQKVVVYDQSSQDVASLSSDCFLTVLLGKLEKSPNSVHL 120
DB 61 DKYLITELIQSHAKHKVIDCSQKVVVYDQSSQDVASLSSDCFLTVLLGKLEKSPNSVHL 120
QY 121 LAGGFA 126
DB 121 LAGADA 126
RESULT 13
Q8ST19 PRELIMINARY; PRT; 657 AA.
ID Q8ST19
AC Q8ST19;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F08B1.1a.
GN F08B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RA Chisoe S.;
RT "The sequence of C. elegans cosmid F08B1.";
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RA Waterston R.;
RT "Direct Submision.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR HSSP: Q16828; IMKP.
DR WormPep: F08B1.1a; CE27918.
DR GO: GO:0016787; F:Hydrolase activity; IEA.
DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO: GO:006470; P:Protein amino acid dephosphorylation; IEA.
DR InterPro: IPR000340; DS phosphatase.
DR InterPro: IPR008343; MAPK phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; Tyr. phosphatase.
DR Pfam: PF00782; DSpC; 1.
DR PRINTS: PRO1764; MAPKPHPTASE.
DR SMART: SM00195; DSpC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANESE_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein; Hydrolase.
SQ SEQUENCE 657 AA; 71002 MW; A0D9153DBE6326B43 CRC64;

Query Match 36.0%; Score 563.5; DB 5; Length 657;
Best Local Similarity 42.0%; Pred. No. 1.7e-44;
Matches 131; Conservative 61; Mismatches 87; Indels 33; Gaps 10;
QY 10 IYTERVALLESSTGKYLIDSRPFVEYNTSHLEAININCSKLMKRLQODRV---LI 65
DB 9 ISTGIALIRAPDPTLVVDCRGFEYNEISHVHSHMNAFFSLIRRLFENLIDNCL 68
QY 66 TELIQSHA-----KHKVIDCSQKVVYDQ-----SSQDVASL-----SSDCFLTVLLGK 110
DB 69 HQLMSSCSGCTKMDKDELDE-----VLVAEBDKPRGNKRIASCPAEPSTAKIMRYLRER 122
QY 111 LEKS--FNSVHLLAGFAFAPFCRPGLCGKS--TLVPCISQPC--PYANIGPRILP 164
DB 123 LEDTDKFRSVVMVEGSGKQFAQOYPCQSSSEGMTPLPSLSQPCLSQPTGD--GITLTP 181
QY 165 NLVAGCQDVNLNKLQNGIGVNLASVYCPKDFIPR--SHPLRPVNDSPCEKILPWL 223
DB 182 NITLGSQIDSLDEMTDADISVINLSMTCPKSCVCKEDKPMRIPVNDSYOKLSPTF 241
QY 224 DKSVDIEKAKASNGCVLVHCLAGISRSATIAIYIMKMDNSLDAVYFVKEKPTISP 283
DB 242 PMAYEFLKCRPRAGKKCLHCLAGISRSPTLAISYIMRYMKGSDDAYRYKRRPSISP 301
QY 284 NRPVPLGQLIDYE 295
DB 302 NPFEMQLLEYE 313
RESULT 14
Q7SZF3 PRELIMINARY; PRT; 367 AA.
ID Q7SZF3
AC Q7SZF3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L.; Feingold E.A.; Grose L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Shennan C.M.; Schuler G.D.;
RA Altschul S.F.; Zeeberg B.; Buettow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;
RA Brownstein M.J.; Udén T.B.; Toshiyuki S.; Carninci P.; Prange C.;
RA Raha S.S.; Loggellano N.A.; Peters G.J.; Abramson R.D.; Mullaney S.J.;
RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;
RA Villalón D.K.; Muny D.M.; Sodeguren E.J.; Lu X.; Gibbs R.A.;
RA Fahney J.; Helton E.; Kettelman M.; Madan A.; Rodrigues S.; Sanchez A.;
RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.; Buttefield Y.S.;
RA Krzywinski M.I.; Skalska U.; Smalins D.E.; Schmech A.; Schein J.E.;
RA Jones S.J.; Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC052477; AAH52477.1; -.

KW Hypothetical protein.
SQ SEQUENCE 367 AA; 40538 MW; F89997A415DAD6AC CRC64;
Query Match 30.2%; Score 472.5; DB 13; Length 367;
Best Local Similarity 34.8%; Pred. No. 3.3e-36;
Matches 106; Conservative 66; Mismatches 110; Indels 19; Gaps 6;
QY 15 LVALLSGTEKVLILDSRPFVEYNTSHLEAININCSKMKRR---LOQDKVLTTELQ 70
DB 15 LKRLMKDDGAKCLLDORSLFASAGHLRGAVVIRCNITVRRRAKSVSLDQILSD-- 71
QY 71 HSAKHKVIDCSQKVVVYDOSQDVASLSDDCFLLVLLGKL-EKSFNS-VHLLAGGFAPF 128
DB 72 DEARSRKSGLSAVILYDERSSDITVMTKDDSTITVHVALCRDFFSTEVYLLKGGYDRF 131
QY 129 SRCPGLC-----EGKSTLVPTCISQPCLPVANIIGPTRILPLVLYGCCQDVINKEL 179
DB 132 STQYPRYCKLTKRTLSTVSSQSSMESCSCATPQHDGGPVELLPFLFGSLALHASKDM 191
QY 180 IQONGIGVYLNASVYCPKDPFLPESHFLRPVNDSPCEKILPWLKSVDFIEKAKASNGC 239
DB 192 LDRMGISALINVSNCPPN-HFEGDYQKCIPEVDNKHEDISSMFLATEFIDSVDSNGR 250
QY 240 VLVHCLAGISRSATIIAYIMKRMDSLDEAYRFYKEKRPITSPNPNFLGQLLDYEKKI 298
DB 251 VLVHCGAGISRSATIIAYIMKRRVRLAEAFEFYKORRSTISPNFSPGQLDPESQV 309
RESULT 15
ID 08R13 PRELIMINARY; PRT; 483 AA.
AC 08R13;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dual specificity phosphatase 10.
GN DUSP10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.
RP [2]
RC STRAIN=C57BL/6J, and MOD;
RC TISSUE=Pancreas, Thymus, and Urinary bladder;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; BC025066; AAH25066.1; -
DR EMBL; AK035293; BAC29019.1; -
DR EMBL; AK050528; BAC34308.1; -
DR EMBL; AK088024; BAC40102.1; -
DR EMBL; AK088186; BAC40196.1; -
DR EMBL; AK088357; BAC40300.1; -
DR HSSP; Q16828; IMKP.
DR MGD; MGI:1927070; Dusp10.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR003440; D: phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PR01764; MAPKPHPTASE.

DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANES_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
SQ SEQUENCE 483 AA; 52532 MW; 7797A1877D986AAB CRC64;
Query Match 30.0%; Score 469; DB 11; Length 483;
Best Local Similarity 35.6%; Pred. No. 1e-35;
Matches 106; Conservative 67; Mismatches 95; Indels 30; Gaps 7;
QY 27 LLDISRFVEYNTSHLEAININCS-KLMKRRLOQDKVLTTELQ-HSAKHKVIDCSQK 84
DB 174 VTIIDCRPFMEYKNSHIOGAVHINCAKISRRRLQGGKITVLDLISCREGKDSFKRIFSKR 233
QY 85 VVVYDOSQDVASLSDDCFLLVLLGKLEKSFNSVHLLAGGFAPFRCFPGLCEGK----- 139
DB 234 IIVYDENTNPSRVTSGQPLHIVLESUKRGEPLVYKGLSFKQNHGRLCDNSIQLOE 293
QY 140 -----STLVPTCISQPCLP-VANIGPTRLIPNLVYGCCQDVINKELIQONGIGY 187
DB 294 CREYGGGASNAASMLPQ--SVPTTPDIEAELTPILPFLFGNQDAODLDTMQRLNIGY 351
QY 188 VLNAS-----YTCPKDPFLPESHFLRPVNDSPCEKILPWLKSVDFIEKAKASNCVY 242
DB 352 VINVTTHLPVLYHEKGLF---NYKRLPATDSKNKRLQYFEEAFEFLEAHQCGKGLIT 407
QY 243 HCLAGISRSATIIAYIMKRMDSLDEAYRFYKEKRPITSPNPNFLGQLLDYEKKIKN 300
DB 408 HCGAGVSRSATIYIAYIMGHTMTMTDAKYPFKGRKPIISPPLNMGQLLEFEEDLNN 465

Search completed: June 21, 2004, 13:22:36
Job time : 24.7353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:10:27 ; Search time 33.7291 Seconds

(without alignments)
2529.847 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302

Perfect score: 1564
Sequence: 1 MAHEMIGTQVTERIVALL.....PNEFGLQLDYKIKKIKQT 302

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1564	100.0	665	5 ABR52381	ABR52381 Protein r
2	1564	100.0	665	5 ABR52407	ABR52407 Protein r
3	1552	99.2	665	4 AAE04834	AAE04834 Human SGP
4	1552	99.2	665	4 AAU09016	AAU09016 Human dua
5	1552	99.2	665	5 AAU79156	AAU79156 Human dua
6	1552	99.2	665	5 AAU09946	AAU09946 Protein s
7	1552	99.2	665	5 AAU75789	AAU75789 Human pro
8	1552	99.2	665	5 ABB97946	ABB97946 Human pro
9	1552	99.2	665	5 AAU79929	AAU79929 Human dua
10	1552	99.2	665	5 ABB97291	ABB97291 Novel hum
11	1545	98.8	665	4 AAB20325	AAB20325 Human pro
12	1545	98.8	665	6 ADA54744	ADA54744 Human pro
13	1544	98.7	665	5 AAU79161	AAU79161 Human dua
14	1544	98.7	665	5 ABR52352	ABR52352 Protein r
15	1544	98.7	672	4 AAM25744	AAM25744 Human pro
16	1542	98.6	302	5 ABR52425	ABR52425 Protein r
17	1542	98.6	665	5 AAU79162	AAU79162 Human dua
18	1526.5	97.6	664	5 ABR52424	ABR52424 Protein r
19	1492	95.4	660	5 ABR52385	ABR52385 Protein r
20	988	63.2	625	5 ABR52382	ABR52382 Protein r
21	988	63.2	625	5 ABR52350	ABR52350 Protein r
22	988	63.0	625	2 AAG73440	AAG73440 Human dua
23	985	63.0	663	2 AAW29150	AAW29150 Dual-spec
24	985	63.0	663	5 ABR52351	ABR52351 Protein r
25	742.5	47.5	375	4 ABG07902	ABG07902 Novel hum

26	741	47.4	140	5 ABR52404	ABR52404 Peptide r
27	716	45.8	140	5 ABR52405	ABR52405 Peptide r
28	715.5	45.7	579	7 ABR08458	ABR08458 Novel pro
29	654.5	41.8	170	4 AAB6436	AAB6436 Human MAP
30	654.5	41.8	170	4 AAE06780	AAE06780 Human dua
31	646	41.3	517	5 AAU79159	AAU79159 Human dua
32	476	30.4	482	4 AAB73225	AAB73225 Human pho
33	469	30.0	444	5 AAO20515	AAO20515 Protein o
34	469	30.0	482	3 AAB18655	AAB18655 A human r
35	469	30.0	482	3 AAB29641	AAB29641 Human dua
36	469	30.0	482	5 ABR52384	ABR52384 Protein r
37	469	29.9	482	5 ABR52383	ABR52383 Protein r
38	467.5	29.9	381	6 ABR52383	ABR52383 Protein r
39	467.5	29.9	381	6 ABR52383	ABR52383 Protein r
40	462.5	29.6	381	7 ABR57087	ABR57087 Mouse MKP
41	461.5	29.5	381	7 ADD48300	ADD48300 Rat Prote
42	461.5	29.5	381	7 ADE62625	ADE62625 Rat Prote
43	440.5	28.2	419	6 ABR55026	ABR55026 Human dua
44	440.5	28.2	419	6 ABR55026	ABR55026 Human dua
45	426	27.2	394	4 AAG67627	AAG67627 Amino aci

ALIGNMENTS

RESULT 1
ID ABR52381 standard; protein; 665 AA.
XX
AC ABR52381;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein relating to the invention SEQ ID NO: 109.
XX
KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;
KW antiproliferative; cardiatic; cytosaric; gene therapy; liver disease;
KW antiproliferative disorder; renal failure; cardiovascular disorder;
KW immunological disorder; arthritis; psoriasis; congenital heart defect;
KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
OS Homo sapiens.
XX
PN WO200257460-A2.
XX
PD 25-JUL-2002.
XX
PF 20-DEC-2001; 2001MO-US050459.
XX
PR 20-DEC-2001; 2000US-0256868P.
PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Jackson DG, Feder J, Nelson T, Mintier G, Ramamathan C, Lee L;
PI Stewers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
PI Kyrestek S, Mcatee P, Suchard S, Banas D;
XX
DR WPI; 2002-599721/64.
XX
DR N-PSDB; ACC60559.
XX
PT Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
PS Claim 5, Fig 12; 801p; English.
XX
CC The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of

CC the invention has antiproliferative, hepatotropic, nephrotropic, CC
CC antirheumatic, antiproliferative, cardiant, and cytostatic activity. The CC
CC polynucleotide may have a use in gene therapy. A polynucleotide or CC
CC polypeptide of the invention is useful for preventing, treating or CC
CC ameliorating a medical condition, e.g. a proliferative disorder. They are CC
CC also useful for treating e.g. liver disease, renal failure, immunological CC
CC disorders including arthritis and psoriasis, cardiovascular disorders CC
CC such as congenital heart defects and congestive heart failure, and CC
CC cancer. A method of the invention is useful for diagnosing a pathological CC
CC condition or susceptibility to a condition in a subject. The present CC
XX sequence 665 AA;

Query Match 100.0%; Score 1564; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 3.4e-164;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVEYNNTSHLEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVEYNNTSHLEAININCSKLMKRRLOQ 60
QY 61 DKVLTITELIQHSAKHKVIDIDCSQKVVVYDQSSQDVASLSDGFLTVLLGKLEKSPNSVHL 120
DB 61 DKVLTITELIQHSAKHKVIDIDCSQKVVVYDQSSQDVASLSDGFLTVLLGKLEKSPNSVHL 120
QY 121 LAGFAEFSPRCFPGCEGKSTLVPTCISOPCLPVANIGTRILPMLYGCORDEVANKELI 180
DB 121 LAGFAEFSPRCFPGCEGKSTLVPTCISOPCLPVANIGTRILPMLYGCORDEVANKELI 180
QY 181 QONGIGYVLANASYTCPKPDFIPESHFLRPVNDSPCEKILPMLDSDVDFIEKAXASNGCV 240
DB 181 QONGIGYVLANASYTCPKPDFIPESHFLRPVNDSPCEKILPMLDSDVDFIEKAXASNGCV 240
QY 241 LVHCLAGISRSATTAIAIYIMKMDMSLDEAYRFVXEKRPPTISPNNFVLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATTAIAIYIMKMDMSLDEAYRFVXEKRPPTISPNNFVLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

RESULT 2
ABR52407
ID ABR52407 standard; protein; 665 AA.

XX ABR52407;

DT 19-JUN-2003 (first entry)

XX Protein relating to the invention SEQ ID NO: 148.

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;
KW antiproliferative; cardiant; cytostatic; gene therapy; liver disease;
KW proliferative disorder; renal failure; cardiovascular disorder;
KW immunological disorder; arthritis; psoriasis; congenital heart defect;
KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.

XX Homo sapiens.
OS
XX
XX WO200257460-A2.
FN
XX
XX 25-JUL-2002.
PD
XX
XX 20-DEC-2001; 2001MO-US050459.
PF
XX
XX 20-DEC-2000; 2000US-0256868P.
PR
XX
XX 30-MAR-2001; 2001US-0280186P.
PR
XX
XX 01-MAY-2001; 2001US-028735P.
PR
XX
XX 05-JUN-2001; 2001US-0295848P.
PR
XX
XX 25-JUN-2001; 2001US-0300465P.

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Jackson DG, Feder J, Nelson T, Mintler G, Ramanathan C, Lee L,
XX Siemers N, Bol D, Schieven G, Finger J, Rodderud CG, Bassolino D,
XX Krystek S, Mcatee P, Suchard S, Bans D,
XX WPI, 2002-599721/64.
DR N-PSDB; ACC60572.

PT Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.

PS Disclosure; Fig 19; 801pp; English.

CC The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic, CC
CC antirheumatic, antiproliferative, cardiant, and cytostatic activity. The CC
CC polynucleotide may have a use in gene therapy. A polynucleotide or CC
CC polypeptide of the invention is useful for preventing, treating or CC
CC ameliorating a medical condition, e.g. a proliferative disorder. They are CC
CC also useful for treating e.g. liver disease, renal failure, immunological CC
CC disorders including arthritis and psoriasis, cardiovascular disorders CC
CC such as congenital heart defects and congestive heart failure, and CC
CC cancer. A method of the invention is useful for diagnosing a pathological CC
CC condition or susceptibility to a condition in a subject. The present CC
XX sequence is used in the exemplification of the invention

XX Sequence 665 AA;

Query Match 100.0%; Score 1564; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 3.4e-164;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVEYNNTSHLEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLESSTGTEKVLIDSRPFVEYNNTSHLEAININCSKLMKRRLOQ 60
QY 61 DKVLTITELIQHSAKHKVIDIDCSQKVVVYDQSSQDVASLSDGFLTVLLGKLEKSPNSVHL 120
DB 61 DKVLTITELIQHSAKHKVIDIDCSQKVVVYDQSSQDVASLSDGFLTVLLGKLEKSPNSVHL 120
QY 121 LAGFAEFSPRCFPGCEGKSTLVPTCISOPCLPVANIGTRILPMLYGCORDEVANKELI 180
DB 121 LAGFAEFSPRCFPGCEGKSTLVPTCISOPCLPVANIGTRILPMLYGCORDEVANKELI 180
QY 181 QONGIGYVLANASYTCPKPDFIPESHFLRPVNDSPCEKILPMLDSDVDFIEKAXASNGCV 240
DB 181 QONGIGYVLANASYTCPKPDFIPESHFLRPVNDSPCEKILPMLDSDVDFIEKAXASNGCV 240
QY 241 LVHCLAGISRSATTAIAIYIMKMDMSLDEAYRFVXEKRPPTISPNNFVLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATTAIAIYIMKMDMSLDEAYRFVXEKRPPTISPNNFVLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

RESULT 3
AAE04834
ID AAE04834 standard; protein; 665 AA.

XX AAE04834;

DT 10-SEP-2001 (first entry)

XX Human GSP002 phosphatase polypeptide.

KW Human; GSP002 phosphatase polypeptide; phosphatase-related disease;
KW immune-related disorder; ocular disease; organ transplant rejection;

	OS	Homo sapiens.	
	XH	Key	Location/Qualifiers
	PH	Domain	1..173
	FT	/label= Catalytic_domain	
	FT	Domain	158..297
	FT	/label= Phosphatase_domain	
	XX	MO200146394-A2.	
	PN		
	PD	28-JUN-2001.	
	XX		
	PF	21-DEC-2000; 2000MO-US034736.	
	XX		
	PR	21-DEC-1999; 99US-017325SP.	
	PR	28-DEC-1999; 99US-0175766P.	
	PR	25-JAN-2000; 2000US-0178078P.	
	PR	31-JAN-2000; 2000US-0179301P.	
	XA	(SUGR-) SUGEN INC.	
	PA		
	XX	Plozman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;	
	PI	Hall RJ, Flanagan P;	
	P1	WPI, 2001-418058/44.	
	DR	N-PDSB; AAD09492.	
	XX		
	PT	Novel phosphatase polypeptide useful for treating cancers, immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases and metabolic disorders.	
	PT		
	PS	Claim 7, Fig 2, 186pp; English.	
	XX		
	CC	The present invention relates to phosphatase polypeptides, nucleotide sequences encoding them, as well as various products and methods useful for the diagnosis and treatment of various phosphatase-related diseases and conditions. Substance that modulates the activity of phosphatase polypeptide is used to treat immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases and CC metabolic disorders, including cancers of tissues, cancers of CC haematopoietic origin, diseases of central and peripheral nervous system, CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic CC lateral sclerosis, viral infections, infections caused by prions, CC bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual CC dysfunction, mood disorder, attention disorders, cognition disorders, CC hypertension, hyperextension, psychotic disorders, neurological disorders, CC dyskinesias and organ transplant rejection. The present amino acid sequence is human SGP002 phosphatase polypeptide. This sequence is CC classified as dual specificity phosphatase (DSP) and MAP kinase CC phosphatase (MKP). SGP002 gene maps to chromosomal position 12p11.1-p12.1 CC XX	
	SQ	Sequence 665 AA;	
		Query Match 99.2%; Score 1552; DB 4; Length 665;	
		Best Local Similarity 99.3%; Pred. No. 7.4e-16;	
		Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
Oy		1 MAHEMIGQIYTERLVALLSGTGVKLLIDSRPFVNTSHLEAININCSGLMKRRLOQ 60	
Dob		1 MAHEMIGQIYTERLVALLSGTGVKLLIDSRPFVNTSHLEAININCSGLMKRRRIQQ 60	
Oy		61 DKVITTEIQHSARKIKVDIDSOKVVVVDDSSGDAVSLSDCFVLVGLKEKSFNVHL 120	

Dd		61	DKVATIELIQSAKHAKVDICSGKVYVVDDSSQDVASLSSDCFLTVLLGLLEKESFNSVHL	120
Oy		121	LAGGFAEFSRCFPGLCEGSKSTLVPTCTISOCLPVAINGPTRILLPNLYLGCGQDVLNKELI	180
Dd		121	LAGEFAEFSRCFPGLCEGSKSTLVPTCTISOCLPVAINGPTRILLPNLYLGCGQDVLNKELM	180
Oy		181	QQNGIGYVLAASYCPEKPDPIPESEHFLRVVNDSFCCKILPMLDKSVDLFEEKASNGCV	240
Dd		181	QQNGIGYVLAASNCPKRPDEIPESHFLRVVNDSFCCKILPMLDKSVDLFEEKASNGCV	240
Oy		241	LVHCLAGISRSATTALAYIMKKMDMSIDEALYRPFYGEKRPTTISPENFLTGOlldYEKKIKN	300
Dd		241	LVHCLAGISRSATTALAYIMKKRDMDSLDEAYRFYKERPTTISPENFLTGOlldYEKKIKN	300
Oy		301	QT 302	
Oy		301	QT 302	
Dd		301	QT 302	
<hr/>				
RESULT 4				
AAU09016	ID	AAU09016	standard; protein; 665 AA.	
XX	AC	AAU09016;		
XX	DT	18-DEC-2001	(first entry)	
XX	DE		Human dual specificity phosphatase 21117.	
XX	KM		Human; dual specificity phosphatase 21117; hepatotropic; cytosstatic;	
KW			hematopoietic disorder; autoimmune disorder; diabetes mellitus;	
KW			rheumatoid arthritis; multiple sclerosis; Crohn's disease;	
KW			liver disorder; erythroïd associated disorder; hemolytic anaemia;	
KW			cellular proliferative; differentiative disorder; leukaemia;	
KW			acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer;	
XX	OS	Homo sapiens.		
XX	FH	Key	Location/Qualifiers	
FT		Domain	11..131	
FT		Domain	/label= Rhodanese-like domain	
FT		Domain	158..297	
FT		Domain	/label= Catalytic domain	
FT		Active-site	/note= "Dual specificity phosphatase catalytic domain"	
FT			242..254	
FT			/label= Tyrosine-specific_protein_phosphatase_active_site	
XX	MO	WO200173059-A2.		
XX	PD	04-OCT-2001.		
XX	PF	23-MAR-2001; 2001WO-US009477.		
XX	PR	24-MAR-2000; 2000US-0191858P.		
XX	PA	(MILL-) MILLENNIUM PHARM INC.		
XX	PL	Meyers RA;		
DR	WP	WI: 2001-611635/70.		
DR	N-PSDB	AAS14639.		
XX	FT	New human dual specificity polypeptides and nucleic acids for diagnosis		
XX	FT	of disease and treatment of e.g. liver disorders.		
XX	PS	Claim 9; Fig 1; 143bp; English.		
CC	XX	The invention relates to two novel human dual specificity phosphataes		
CC	XX	designated 21117 and 38692, the nucleic acids encoding them (including		
CC	XX	fragments, allelic variants, their complements or nucleic acids that		
CC	XX	hybridise to them) and antibodies raised against the proteins. The		

CC antibody is useful for detecting the presence of the polypeptide, and the
CC nucleic acid fragments are useful for detecting the presence of the
CC nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
CC antisense sequences) are useful for modulating the activity or expression
CC of the polypeptide/nucleic acid, useful for the treatment of e.g.
CC hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
CC listed in the specification) liver disorders, erythroid associated
CC disorders (e.g. haemolytic anaemia) cellular proliferative or
CC differentiative disorders, leukaemias (e.g. acute myeloid leukaemia),
CC carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and
CC 38692 are also useful for modulating the proliferation, survival,
CC migration or differentiation of a 38692 or 21117-expressing cell. The
CC polypeptide and nucleic acids are useful for identifying modulating
CC agents. The present sequence represents the dual specificity phosphatase
CC 21117

XX Sequence 665 AA;

Query Match 99.2%; Score 1552; DB 4; Length 665;
Best Local Similarity 99.3%; Pred. No. 7,4e-163; Indels 0; Gaps 0;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALLSEGTKEKVLIDSRPFVEYNTSHLEAININCSKMKRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLSEGTKEKVLIDSRPFVEYNTSHLEAININCSKMKRRLOQ 60
QY 61 DKVLTTELIOHSAKHKVDIDCSQKVVVYDQSSQDVASISDCEFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLTTELIOHSAKHKVDIDCSQKVVVYDQSSQDVASISDCEFLTVLLGKLEKSFNSVHL 120
QY 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTPTILPNLYLGCQRDVLANKELI 180
DB 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTPTILPNLYLGCQRDVLANKELI 180
QY 181 OONGIGYVLNASYTCRKPDPFIPESHFLRPVNDSEFCETILPMDKSDVDFEKAASNGCV 240
DB 181 OONGIGYVLNASYTCRKPDPFIPESHFLRPVNDSEFCETILPMDKSDVDFEKAASNGCV 240
QY 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

RESULT 5
AAU79156
ID AAU79156 standard; protein; 665 AA.

XX AAU79156;

XX 02-JUL-2002 (first entry)

XX Human dual-specificity phosphatase-3 (DSP-16) protein.
XX Human, dual-specificity phosphatase-3; DSP-16; autoimmune disease;
XX mtogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
XX cancer; graft-versus-host disease; allergy; metabolic disease;
XX abnormal cell growth; abnormal cell proliferation; contact inhibition;
XX cell cycle abnormality; anchorage independent cell growth; apoptosis;
XX intercellular adhesion; DSP-16 modulator; chromosome 12p.

XX Homo sapiens.

XX Key Location/Qualifiers
XX Domain 242..251
XX /label=Active_site_domain

XX MO200226997-A2.

PD 04-APR-2002.
XX
XX 25-SEP-2001; 2001WO-US030124.
XX
XX 26-SEP-2000; 2000US-0235487P.
XX
XX (CEPT-1) CEPTYR INC.
XX
XX Luche RM, Wei B;
XX WPI; 2002-315802/35.
XX N-PSDB; ABK4759C.
XX
XX New DSP-16 polypeptide, useful for identifying modulators of its
XX activity, which can be used in the treatment of disorders such as
XX Duchenne muscular dystrophy, or cancer.

PS Claim 1; Fig 2; 87pp; English.

CC The present invention relates to a new polypeptide, DSP-16, having a 665
CC amino acid sequence, given in the specification, or a variant having at
CC least 50 % identical residues, which retains the ability to
CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
CC invention can be used for identifying agents which modulate DSP-16
CC activity, for modulation of a proliferative response in a cell, survival
CC of a cell, or differentiation of a cell. The cell displays contact
CC inhibition of cell growth or anchorage independent growth and may display
CC altered intercellular adhesion. The agent may modulate apoptosis, or the
CC cell cycle. The identified modulators can be used to treat Duchenne
CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
CC cell proliferation, and cell cycle abnormalities. The present amino acid
CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
CC protein of the invention. This sequence is encoded by the human DSP-16
CC gene located on chromosome 12p

XX Sequence 665 AA;

Query Match 99.2%; Score 1552; DB 5; Length 665;
Best Local Similarity 99.3%; Pred. No. 7,4e-163; Indels 0; Gaps 0;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQVTERLVALLSEGTKEKVLIDSRPFVEYNTSHLEAININCSKMKRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLSEGTKEKVLIDSRPFVEYNTSHLEAININCSKMKRRLOQ 60
QY 61 DKVLTTELIOHSAKHKVDIDCSQKVVVYDQSSQDVASISDCEFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLTTELIOHSAKHKVDIDCSQKVVVYDQSSQDVASISDCEFLTVLLGKLEKSFNSVHL 120
QY 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTPTILPNLYLGCQRDVLANKELI 180
DB 121 LAGFAEFSRCPGPGCEGKSTLVPTCISQPCLPVANIPTPTILPNLYLGCQRDVLANKELI 180
QY 181 OONGIGYVLNASYTCRKPDPFIPESHFLRPVNDSEFCETILPMDKSDVDFEKAASNGCV 240
DB 181 OONGIGYVLNASYTCRKPDPFIPESHFLRPVNDSEFCETILPMDKSDVDFEKAASNGCV 240
QY 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATTAIAIYIMKRMDSIDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

RESULT 6
AAU09946
ID AAU09946 standard; protein; 665 AA.
XX
XX AAU09946;

XT	18-JUN-2002	(first entry)
DE	Protein sequence of human (dual specificity phosphatase) DUSP-10.	
XX	Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke;	
KM	neuronal degeneration syndrome; Alzheimer's disease; depression;	
KM	schizophrenia; asthma; immune disorder; inflammatory process; arthritis;	
KM	osteoporosis; diabetes.	
OS	Homo sapiens.	
XX	WO2001.77340-A1.	
PN	18-OCT-2001.	
PD	06-APR-2001; 2001WO-EP003966.	
XX	10-APR-2000; 2000EP-00107143.	
PR	(MERE) MERCK PATENT GMBH.	
XX	Duecker K;	
PA	WI: 2002-010917/01.	
PI	N-PSDB; AAS15768.	
DR	Novel dual specificity phosphatase polypeptides useful for treating	
PT	cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's	
PT	disease, depression, schizophrenia, asthma and immune disorders.	
PS	Claim 2; Page 37-39; 43pp; English.	
XX	The present invention relates to a new isolated dual specificity	
CC	phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid	
CC	sequence that is fully defined in the specification. The invention also	
CC	provides a sequence encoded by a 3059 nucleotide sequence fully defined	
CC	in the specification, and a sequence having at least 95 % identity to the	
CC	polypeptide, or fragments or variants of DUSP-10. The invention is useful	
CC	for treating cancer e.g. leukaemia, colon carcinoma, lung cancer,	
CC	prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,	
CC	stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,	
CC	schizophrenia, cardiac myotrophies, asthma, immune disorders,	
CC	inflammatory processes e.g. arthritis, bowel disease, type 1 diabetes,	
CC	osteoporosis, diabetes and diabetes associated diseases. The molecules of	
CC	the invention are also useful as vaccines for inducing immunological	
CC	response in a mammal, in disease diagnosis and in assays for screening	
CC	agonistic or antagonistic compounds. Other uses of the invention include	
CC	identifying membrane bound or soluble receptors, as a diagnostic reagent,	
CC	in chromosome localisation studies, and as a valuable tool in tissue	
CC	expression studies. The present sequence represents the dual specificity	
CC	phosphatase, DUSP-10, protein of the invention	
XX		
SO	Sequence 665 AA:	
QY	Query Match 99.2%; Score 1552; DB 5; Length 665;	
QY	Best Local Similarity 99.3%; Pred. No. 7.4e-163;	
QY	Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
Db	1 MAHEMIGTQYITERLVALLSEGTENVLLIDSRPFYENTYSHLEAININCSKLMRRQLQ 60	
QY	1 MAHEMIGTQYITERLVALLSEGTENVLLIDSRPFYENTYSHLEAININCSKLMRRQLQ 60	
Db	61 DKVLTITELIGHSAKKHVDIDCSQKVVYDQSSQDVASLSSDCEFLVLLGKLEKSFNSVHL 120	
QY	61 DKVLTITELIGHSAKKHVDIDCSQKVVYDQSSQDVASLSSDCEFLVLLGKLEKSFNSVHL 120	
Db	121 LAGGAFAEBSRCFPGICEKSTLVPTCISQPCLPVANIIGPTRILPNLYLGGQDVAINKELI 180	
QY	121 LAGGAFAEBSRCFPGICEKSTLVPTCISQPCLPVANIIGPTRILPNLYLGGQDVAINKELI 180	
Db	121 LAGGAFAEBSRCFPGICEKSTLVPTCISQPCLPVANIIGPTRILPNLYLGGQDVAINKELI 180	
QY	181 QONGIGYLVANSYTPKPDFIPESHFLAVPVNDSCEKILPMVLDNSVDFIEAKASNGCV 240	
Db	181 QONGIGYLVANSYTPKPDFIPESHFLAVPVNDSCEKILPMVLDNSVDFIEAKASNGCV 240	

QY	241	LVHCLAGISRSRAATTAIAIYIMRMDMSLDEAYRFEVKEKPTTISPNNFGLQLDPEKKIKN	300
Db	241	LVHCLAGISRSRAATTAIAIYIMRMDMSLDEAYRFEVKEKPTTISPNNFGLQLDPEKKIKN	300
QY	301	QT 302	
Db	301	QT 302	
RESULT 7			
ID	AAU75789	standard; protein; 665 AA.	
XX	AAU75789		
AC	AAU75789		
DT	08-MAY-2002	(first entry)	
XX			
DE	Human protein phosphatase 7 (PP7) protein sequence.		
XX			
KM	Human; protein phosphatase 7; PP7; immune system disorder; AIDS;		
KM	acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;		
KM	Crohn's disease; neurological disorder; epilepsy; Huntington's disease;		
KM	dementia; Parkinson's disease; developmental disorder; Down's syndrome;		
KM	cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;		
KM	melanoma; myeloma sarcoma.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Domain	11..131	
FT		/label= Rhodanese_like_domain	
FT	Region	15..170	
FT		/note= "Dual specificity protein phosphatase"	
FT	Region	85..298	
FT		/note= "VH1-type dual specificity protein phosphatase"	
FT	Domain	158..297	
FT		/label= Catalytic domain DSPC	
FT		/note= "Dual specificity phosphatase"	
FT	Region	220..281	
FT		/note= "Tyrosine specific protein phosphatases active	
FT	Region	site tyr phosphatase.prf"	
FT		242..254	
FT		/note= "Tyrosine phosphatase"	
FT	Region	242..252	
FT		/note= "Tyrosine specific protein phosphatases"	
XX			
FN	WO200210363-A2.		
XX			
PD	07-FEB-2002.		
XX			
PF	26-JUL-2001; 2001WO-US023716.		
XX			
PR	28-JUL-2000; 2000US-0221679P.		
PR	03-AUG-2000; 2000US-0223272P.		
PR	10-AUG-2000; 2000US-0224309P.		
PR	18-AUG-2000; 2000US-0226728P.		
PR	30-AUG-2000; 2000US-0228254P.		
PR	08-SEP-2000; 2000US-0231366P.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
PI	Tang YF, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;		
PI	Stewart EA, Gandhi AR, Patterson C, Lee EA, Hafalla AJA, Lu DAM;		
PI	Tribouley CM, Griffin JA, Baughn MK, Yue H, Warren BA, Nguyen DB;		
XX	Walia NK, Kearney L;		
XX	WPI; 2002-188735/24.		
DR	N-Psdb; ABK14474.		
XX			
PT	New protein phosphatases, useful for diagnosing, treating or preventing		
PT	immune system disorders (e.g. Crohn's disease), neurological disorders		
PT	(e.g. Parkinson's disease), or cell proliferative disorders (e.g.		

PT cancers).

XX Claim 1; Page 107-108; 117pp; English.

CC The present invention relates to a new polypeptide, a naturally occurring
CC amino acid sequence at least 95 % identical to it, a biologically active
CC fragment of it or an immunogenic fragment of it. The polypeptides,
CC polynucleotides, agonists and antagonists are useful for diagnosing,
CC treating or preventing disorders associated with aberrant expression of
CC protein phosphatases (PP), particularly immune system disorders e.g.
CC acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
CC asthma or Crohn's disease, neurological disorders e.g. epilepsy,
CC Huntington's disease, dementia or Parkinson's disease, developmental
CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
CC or sarcoma. The present amino acid sequence represents human protein
CC phosphatase 7 (PP7) which is one of several human protein phosphatases
CC (AAU7583-AAU7592) of the invention

SQ Sequence 665 AA;

Query Match 99.2%; Score 1552; DB 5; Length 665;
Best Local Similarity 99.3%; Pred. No. 7.4e-163;

Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MAHEMIGTQIVTERLVALLSEGTETKLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
DB 1 MAHEMIGTQIVTERLVALLSEGTETKLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
QY 61 DKVLTITELIOHSAKHKVIDCSQKVVYDQSSQDVASISDPCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLTITELIOHSAKHKVIDCSQKVVYDQSSQDVASISDPCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCPGICBCKSTLPTCTISQPCLPVANTGPTRIIPNLVYLGCGQDVNLKELI 180
DB 121 LAGGFAEFSRCPGICBCKSTLPTCTISQPCLPVANTGPTRIIPNLVYLGCGQDVNLKELI 180
QY 181 QONGIGYVLNASYTCRPFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGYVLNASYTCRPFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIYIMKMDMSLDEAYRFVEKRPPTISPNFNFLQGLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIYIMKMDMSLDEAYRFVEKRPPTISPNFNFLQGLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

```

RESULT 8

AB97946 ID ABB97946 standard; protein; 665 AA.

AC ABB97946;

DT 06-SEP-2002 (first entry)

DE Human protein sequence #13.

KW Human; brain; tonsil; hippocampus; foetal brain; diagnosis.

OS Homo sapiens.

PN WO200252005-A1.

PD 04-JUL-2002.

PF 20-DEC-2001; 2001WO-JP011217.

PR 22-DEC-2000; 2000JP-00389742.

PA (KAZU-) KAZUSA DNA RES INST FOUND.

PA (CELE-) CELESTAR LEXICO-SCI LTD.

PI Ohara O, Nagase T, Nakajima D;

DR WPI; 2002-500762/53.

DR N-PSDB; ABBN83966.

PT Genes and their expression products cloned from human cDNA libraries for
PT treatment and diagnosis of diseases associated with their expression.

PS Claim 1(a); Page 112-116; 238pp; Japanese.

CC The invention relates to DNA encoding polypeptides directly cloned from
CC cDNA libraries originating in adult whole brain, human tonsil, human
CC adult hippocampus and human foetal whole brain. Polypeptides and
CC polynucleotides of the invention may be used in the investigation of
CC differential expression of the DNA sequences in normal subjects and
CC disease patients. They may also be used in the production of antibodies
CC oligonucleotide probes and DNA chips for diagnosis and identification of
CC drugs for treatment of diseases with which the DNA sequences are
CC associated. The sequences given in records ABB97934-ABB97964 represent
CC human proteins of the invention

SQ Sequence 665 AA;

Query Match 99.2%; Score 1552; DB 5; Length 665;
Best Local Similarity 99.3%; Pred. No. 7.4e-163;

Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MAHEMIGTQIVTERLVALLSEGTETKLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
DB 1 MAHEMIGTQIVTERLVALLSEGTETKLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
QY 61 DKVLTITELIOHSAKHKVIDCSQKVVYDQSSQDVASISDPCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLTITELIOHSAKHKVIDCSQKVVYDQSSQDVASISDPCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCPGICBCKSTLPTCTISQPCLPVANTGPTRIIPNLVYLGCGQDVNLKELI 180
DB 121 LAGGFAEFSRCPGICBCKSTLPTCTISQPCLPVANTGPTRIIPNLVYLGCGQDVNLKELI 180
QY 181 QONGIGYVLNASYTCRPFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGYVLNASYTCRPFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIYIMKMDMSLDEAYRFVEKRPPTISPNFNFLQGLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIYIMKMDMSLDEAYRFVEKRPPTISPNFNFLQGLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

```

RESULT 9

AAU79929 ID AAU79929 standard; protein; 665 AA.

AC AAU79929;

DT 02-JUL-2002 (first entry)

DE Human dual specificity phosphatase 21117 protein.

KW Human; dual specificity phosphatase 21117; erythroid-related disorder;

KW haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;

OS erythrocytosis; liver-related disorder; cancer.

PA Homo sapiens.

Key Location/Qualifiers
FH 11.131
FT /label= Rhodanese_1like_domain

FT Region 21..24
 FT /note="Casein kinase II phosphorylation site"
 FT 91..94
 FT /note="Casein kinase II phosphorylation site"
 FT 158..297
 FT /note="Dual specificity phosphatase catalytic domain"
 FT 214..217
 FT /note="Casein kinase II phosphorylation site"
 FT 242..254
 FT /note="Tyrosine specific protein phosphatase active site"
 FT 242..254
 FT /note="C-X5-R motif"
 FT 266..269
 FT /note="Casein kinase II phosphorylation site"
 FT 369..372
 FT /note="Casein kinase II phosphorylation site"
 FT 421..424
 FT /note="Casein kinase II phosphorylation site"
 FT 434..437
 FT /note="Casein kinase II phosphorylation site"
 FT 458..461
 FT /note="Casein kinase II phosphorylation site"
 FT 508..511
 FT /note="Casein kinase II phosphorylation site"
 FT 589..592
 FT /note="Casein kinase II phosphorylation site"
 FT 612..615
 FT /note="Casein kinase II phosphorylation site"
 FT 617..620
 FT /note="Casein kinase II phosphorylation site"
 FT 642..645
 FT /note="Casein kinase II phosphorylation site"
 FT Region
 PN US002034807-A1.
 XX 21-MAR-2002.
 XX 23-MAR-2001; 2001US-00816494.
 XX 24-MAR-2000; 2000US-0191858P.
 XX (MEYE/) MEYERS R A.
 XX Meyers RA;
 XX WPI; 2002-351088/38.
 XX N-PSDB; ABK49402.
 XX
 PT New nucleic acids, designated 38692 and 21117, encoding dual specificity
 PT phosphatases for treating cell proliferation and differentiation
 PT disorders including hematopoietic and erythroid-related disorders and
 PT cancers.
 PT
 PS Claim 8; Fig 1; 76pp; English.
 XX
 CC The present invention relates to new nucleic acids designated 38692 and
 CC 21117 encoding dual specificity phosphatase family members. The nucleic
 CC acid, polypeptide encoded by it, and antibody specific for the
 CC polypeptide may be used to diagnose and treat haematopoietic-related
 CC disorders such as leukaemias and autoimmune diseases, erythroid-related
 CC disorders such as anaemias and erythrocytosis, liver-related disorders,
 CC and cancers, particularly of the breast, colon, adipose, prostate and
 CC lung. The present amino acid sequence represents the human dual
 CC specificity phosphatase 21117 protein of the invention, as described
 CC above
 XX
 SO Sequence 665 AA:
 Query Match 99.2%; Score 1552; DB 5; Length 665;
 Best Local Similarity 99.3%; Pred. No. 7,4e-163;
 Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRLQQ 60
 DB 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRLQQ 60
 QY 61 DKVLITELIQSHAKRVIDCSQKVVVYDQSDVASLSDDCLTVLLGKLEKSFNSVHL 120
 DB 61 DKVLITELIQSHAKRVIDCSQKVVVYDQSDVASLSDDCLTVLLGKLEKSFNSVHL 120
 QY 121 LAGGFAPSRCPGCGESTLVPCTISQPCPVANIGPRLIPNLYLGGQRLVNLKELI 180
 DB 121 LAGGFAPSRCPGCGESTLVPCTISQPCPVANIGPRLIPNLYLGGQRLVNLKELI 180
 QY 181 QONGIGYVNASYTCRPPDIPESHFLRVVNDSPCEKILPMDKSDVFIEKAKANGCV 240
 DB 181 QONGIGYVNASYTCRPPDIPESHFLRVVNDSPCEKILPMDKSDVFIEKAKANGCV 240
 QY 241 LVHCLAGISRSATIAIYIMKMDMSLDEAYRFEKERTISPNFPLGQLDYEKKIKN 300
 DB 241 LVHCLAGISRSATIAIYIMKMDMSLDEAYRFEKERTISPNFPLGQLDYEKKIKN 300
 QY 301 QT 302
 DB 301 QT 302
 DB 301 QT 302
 RESULT 10
 ABB97291
 ID ABB97291 standard; protein; 665 AA.
 XX
 AC ABB97291;
 XX
 DT 27-JUN-2002 (first entry)
 XX
 DE Novel human protein SEQ ID NO: 559.
 XX
 XX Human; anti-nausea; vulnary; anti-inflammatory; immunomodulator;
 KM antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KM neuroprotective; antiparkinsonian; protein therapy; EST;
 XX expressed sequence tag.
 XX
 OS Homo sapiens.
 OS
 PN MO200222660-A2.
 XX
 PD 21-MAR-2002.
 XX
 PF 10-SEP-2001; 2001WO-US026015.
 XX
 PR 11-SEP-2000; 2000US-00659671.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao Q, Ren F;
 PI Xue AJ, Yang Y, Weinman T, Dermanac RT;
 XX
 DR N-PSDB; ABN32477.
 DR
 XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.
 PT
 PS Example 2; SEQ ID NO 559; 509pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention

XX Sequence 665 AA;
SQ
Query Match 99.2%; Score 1552; DB 5; Length 665;
Best Local Similarity 99.3%; Pred. No. 7,46-163;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAHEMIGTQIVTERLVALLSEGTETKVLIDSRPEVNTSHILEANINOSKLMKRLQ 60
DB 1 MAHEMIGTQIVTERLVALLSEGTETKVLIDSRPEVNTSHILEANINOSKLMKRLQ 60
QY 61 DKVLITLILQHSARKVDIDCSQVTVVYDSSQDVASLSSDCELTVLLKLEKSPNSVHL 120
DB 61 DKVLITLILQHSARKVDIDCSQVTVVYDSSQDVASLSSDCELTVLLKLEKSPNSVHL 120
QY 121 LAGGFARFRCFPGLCGKSTLVPTCISQPCLVANIGPRLIPNLVLCQDVLNKLKI 180
DB 121 LAGGFARFRCFPGLCGKSTLVPTCISQPCLVANIGPRLIPNLVLCQDVLNKLKI 180
QY 181 QONGIGVYNASTCTCPKDFIPESHFLRVVNDSPCEKILPMLDKSVDFIERAKASNGCV 240
DB 181 QONGIGVYNASTCTCPKDFIPESHFLRVVNDSPCEKILPMLDKSVDFIERAKASNGCV 240
QY 241 LVHCLAGISRSATITAIYIMKRWMSLDRAVRFVKEKRTIPNRPFLGQLIDYEKKIKN 300
DB 241 LVHCLAGISRSATITAIYIMKRWMSLDRAVRFVKEKRTIPNRPFLGQLIDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
RESULT 11
AAB20325
ID AAB20325 standard; protein; 666 AA.
AC AAB20325;
XX
DT 29-MAY-2001 (first entry)
XX
DE Human protein phosphatase and kinase protein-4.
XX
KW Protein phosphatase and kinase protein; PPHKP-4; human;
KW gastrointestinal disorder; immune system disorder; neurological disorder;
XX cell proliferative disorder; cancer; diagnosis; therapy.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 12 /note= "O-phosphorylated"
FT Modified-site 21 /note= "O-phosphorylated"
FT Modified-site 23 /note= "O-phosphorylated"
FT Modified-site 38 /note= "O-phosphorylated"
FT Modified-site 38 /note= "N-glycosylated"
FT Modified-site 38 /note= "N-glycosylated"
FT Modified-site 38 /note= "N-glycosylated"
FT Modified-site 49 /note= "N-glycosylated"
FT Modified-site 72 /note= "N-glycosylated"
FT Modified-site 82 /note= "O-phosphorylated"
FT Modified-site 82 /note= "O-phosphorylated"
FT Region 85..298 /note= "VHL-type dual specificity phosphatase signature"
FT Modified-site 91 /note= "O-phosphorylated"
FT Modified-site 190 /note= "O-phosphorylated"
FT Modified-site 212 /note= "N-glycosylated"
FT Modified-site 212 /note= "N-glycosylated"

FT Modified-site 214 /note= "O-phosphorylated"
FT Active-site 220..280 /note= "tyrosine specific protein phosphatase"
FT Region 237..278 /note= "Y phosphatase signature"
FT Modified-site 266 /note= "O-phosphorylated"
FT Modified-site 280 /note= "O-phosphorylated"
FT Modified-site 300 /note= "O-phosphorylated"
FT Modified-site 369 /note= "N-glycosylated"
FT Modified-site 393 /note= "O-phosphorylated"
FT Modified-site 421 /note= "O-phosphorylated"
FT Modified-site 422 /note= "O-phosphorylated"
FT Modified-site 434 /note= "O-phosphorylated"
FT Modified-site 439 /note= "O-phosphorylated"
FT Modified-site 468 /note= "O-phosphorylated"
FT Modified-site 471 /note= "O-phosphorylated"
FT Modified-site 479 /note= "O-phosphorylated"
FT Modified-site 528 /note= "O-phosphorylated"
FT Modified-site 590 /note= "O-phosphorylated"
FT Modified-site 597 /note= "O-phosphorylated"
FT Modified-site 605 /note= "O-phosphorylated"
FT Modified-site 610 /note= "O-phosphorylated"
FT Modified-site 613 /note= "O-phosphorylated"
FT Modified-site 618 /note= "O-phosphorylated"
FT Modified-site 628 /note= "O-phosphorylated"
FT Modified-site 641 /note= "O-phosphorylated"
FT Modified-site 643 /note= "N-glycosylated"
FT Modified-site 643 /note= "O-phosphorylated"
XX
XX WO200120004-A2.
XX
XX 22-MAR-2001.
XX
XX 14-SEP-2000; 2000WO-US025515.
XX
XX 15-SEP-1999; 99US-015414P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y,
XX Lu DM;
XX
XX WPI; 2001-244811/25.
XX
XX DR N-PSDB; AAF30479.
XX
XX Novel human protein phosphatase and kinase proteins for diagnosis,
XX treatment and prevention of gastrointestinal, immune system, neurological
XX and cell proliferative disorders.
XX
XX Claim 1; Page 87-88; 103pp; English.
XX

CC The present sequence is that of novel human protein phosphatase and
 CC kinase protein PPHK-4, as predicted from Incyte Clone ID No. 1234795CB1
 CC (see AA030479). Tissues that express PPHK-4 (as a fraction of total
 CC tissues expressing PPHK-4) include gastrointestinal (0.365),
 CC cardiovascular (0.154), nervous (0.154) and reproductive (0.154).
 CC Diseases or conditions associated with tissues expressing PPHK-4 (as a
 CC fraction of total tissues expressing PPHK-4) include cancer (0.692),
 CC inflammation or trauma (0.308) and cell proliferation (0.231). The
 CC encoded protein shows homology to mouse neuronal tyrosine threonine
 CC phosphatase 1. The invention provides human PPHK-1 to -11 polypeptides
 CC (see AA020322-32) and polynucleotides (see AA030476-86). It also provides
 CC expression vectors, host cells, antibodies, agonists and antagonists, as
 CC well as methods for diagnosing, treating or preventing disorders
 CC associated with expression of PPHK, including gastrointestinal
 CC disorders, immune system disorders, neurological disorders and cell
 CC proliferative disorders, including cancer

CC Sequence 666 AA;

Query Match 99.2%; Score 1552; DB 4; Length 666;
 Best Local Similarity 99.3%; Pred. No. 7.4e-163; Indels 0; Gaps 0;
 Matches 300; Conservative 1; Mismatches 1;

QY 1 MAHEMIGTQIVTERLVALLSGTEKYLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
 DB 1 MAHEMIGTQIVTERLVALLSGTEKYLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
 QY 61 DKVLITELIQHSARKKVIDCSQKVVYVYDSSQDVASLSDDCFLYLLGKLEKSFNSVHL 120
 DB 61 DKVLITELIQHSARKKVIDCSQKVVYVYDSSQDVASLSDDCFLYLLGKLEKSFNSVHL 120
 QY 121 LAGGFAPSRCPGCLCEGKSTLVPTCISOPCLPVANIGFTRILPNLYLGGQRDVNLKELI 180
 DB 121 LAGGFAPSRCPGCLCEGKSTLVPTCISOPCLPVANIGFTRILPNLYLGGQRDVNLKELI 180
 QY 181 QONGIGYVNAASYTCRPPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 DB 181 QONGIGYVNAASYTCRPPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 QY 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPPTISPENFLGQLDYEKKIKN 300
 DB 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPPTISPENFLGQLDYEKKIKN 300
 QY 301 QT 302
 DB 301 QT 302

RESULT 12

ADA54744
 ID ADA54744 standard; protein; 665 AA.

AC ADA54744;

XX 20-NOV-2003 (first entry)

XX Human protein, SEQ ID 2312.

XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 XX Gene Therapy; human; secretory protein; membrane proteins; cancer;
 XX inflammatory disease; osteoporosis; neurological disease.

XX Homo sapiens.

XX Homo sapiens.

XX EP1293569-A2.

XX 19-MAR-2003.

XX 21-MAR-2002; 2002EP-0006586.

XX 14-SEP-2001; 2001JP-0032831.

XX 24-JAN-2002; 2002US-0350435P.

PA (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
 PI Yamamoto U, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y;

DR WPI: 2003-395539/38.
 DR N-PSDB; ADA53105.

XX New polynucleotides encoding full-length polypeptides, e.g. secretory
 XX PT and/or membrane proteins, useful for developing medicines for diseases in
 XX PT which the gene is involved, or as target molecules for gene therapy.

XX Claim 14; SEQ ID NO 2312; 205pp; English.

CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA5433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.

XX Sequence 665 AA;

Query Match 98.8%; Score 1545; DB 6; Length 665;
 Best Local Similarity 99.0%; Pred. No. 4.4e-162; Indels 0; Gaps 0;
 Matches 299; Conservative 1; Mismatches 2;

QY 1 MAHEMIGTQIVTERLVALLSGTEKYLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
 DB 1 MAHEMIGTQIVTERLVALLSGTEKYLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
 QY 61 DKVLITELIQHSARKKVIDCSQKVVYVYDSSQDVASLSDDCFLYLLGKLEKSFNSVHL 120
 DB 61 DKVLITELIQHSARKKVIDCSQKVVYVYDSSQDVASLSDDCFLYLLGKLEKSFNSVHL 120
 QY 121 LAGGFAPSRCPGCLCEGKSTLVPTCISOPCLPVANIGFTRILPNLYLGGQRDVNLKELI 180
 DB 121 LAGGFAPSRCPGCLCEGKSTLVPTCISOPCLPVANIGFTRILPNLYLGGQRDVNLKELI 180
 QY 181 QONGIGYVNAASYTCRPPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 DB 181 QONGIGYVNAASYTCRPPDFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKANGCV 240
 QY 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPPTISPENFLGQLDYEKKIKN 300
 DB 241 LVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKRPPTISPENFLGQLDYEKKIKN 300
 QY 301 QT 302
 DB 301 QT 302

RESULT 13

AAU79161
 ID AAU79161 standard; protein; 665 AA.

XX AAU79161;

XX 02-JUL-2002 (first entry)

XX Human dual-specificity phosphatase-3 (DSP-16) mutant protein #1.

XX Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 XX mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 XX cancer; graft-versus-host disease; allergy; metabolic disease;
 XX abnormal cell growth; abnormal cell proliferation; contact inhibition;
 XX cell cycle; abnormality; anchorage independent cell growth; apoptosis;
 XX intercellular adhesion; DSP-16 modulator; mutant; mutein.

XX Homo sapiens.

XX Homo sapiens.

XX Synthetic.

XX Synthetic.


```

FH Key Location/Qualifiers
FT Misc-difference 213 /note= "Wild-type Asp substituted by Ala"
XX
XX PN WO200226997-A2.
XX PD 04-APR-2002.
XX
XX 25-SEP-2001; 2001WO-US030124.
XX
XX 26-SEP-2000; 2000US-0235487P.
XX
XX (CEPT-) CEPTYR INC.
XX
XX Luche RM, Wei B;
XX
XX WPI; 2002-315802/35.
XX
XX New DSP-16 polypeptide, useful for identifying modulators of its
XX activity, which can be used in the treatment of disorders such as
XX Duchenne muscular dystrophy, or cancer.
XX
XX Claim 46; Page; 87pp; English.
XX
XX The present invention relates to a new polypeptide, DSP-16, having a 665
XX amino acid sequence, given in the specification, or a variant having at
XX least 50 % identical residues, which retains the ability to
XX dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
XX invention can be used for identifying agents which modulate DSP-16
XX activity, for modulation of a proliferative response in a cell, survival
XX of a cell, or differentiation of a cell. The cell displays contact
XX inhibition of cell growth or anchorage independent growth and may display
XX altered intercellular adhesion. The agent may modulate apoptosis, or the
XX cell cycle. The identified modulators can be used to treat Duchenne
XX muscular dystrophy, cancer, graft-versus-host disease, autoimmune
XX diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
XX cell proliferation, and cell cycle abnormalities. The present amino acid
XX sequence represents the human dual-specificity phosphatase-3 (DSP-16)
XX mutant protein #1. Note: This sequence is not shown in the specification
XX but is derived from the wild-type human DSP-16 (AAU79156) protein given
XX in Figure 2 of the specification
XX
XX SQ Sequence 665 AA;
XX
XX Query Match 98.7%; Score 1544; DB 5; Length 665;
XX Best Local Similarity 99.0%; Pred. No. 5.7e-162;
XX Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNTSHIILPAININCSKLMKRRLQ 60
XX |||||||
XX DB 1 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNTSHIILPAININCSKLMKRRLQ 60
XX
XX QY 61 DKVLTIELIOHSAKHKVIDCSQKVVVYDOSQDVASLSDFLTVLGKLEKSFNSVHL 120
XX |||||||
XX DB 61 DKVLTIELIOHSAKHKVIDCSQKVVVYDOSQDVASLSDFLTVLGKLEKSFNSVHL 120
XX
XX QY 121 LAGFAFSPSCFPGICSGKSTLVPTCISQCLPANIPTIILPNLVILGCGORDVLANEIM 180
XX |||||||
XX DB 121 LAGFAFSPSCFPGICSGKSTLVPTCISQCLPANIPTIILPNLVILGCGORDVLANEIM 180
XX
XX QY 181 QONGIGVVLNASTYCPKRPDIIPESHFLRPVNDSCFCKILPMDKSVDFIKAVANGCV 240
XX |||||||
XX DB 181 QONGIGVVLNASTYCPKRPDIIPESHFLRPVNDSCFCKILPMDKSVDFIKAVANGCV 240
XX
XX QY 241 LVHCLAGISRSATIAIAYIMKRMMSIDEAVRFYKPKPTISPNPFLGQLIDYEKKIKN 300
XX |||||||
XX DB 241 LVHCLAGISRSATIAIAYIMKRMMSIDEAVRFYKPKPTISPNPFLGQLIDYEKKIKN 300
XX
XX QY 301 QT 302
XX ||
XX DB 301 QT 302

```

```

RESULT 14
ABR52352
ID ABR52352 standard; protein; 665 AA.
XX
XX AC ABR52352;
XX
XX DT 19-JUN-2003 (first entry)
XX
XX DE Protein relating to the invention SEQ ID NO: 42.
XX
XX KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;
XX antiproliferative; cardiac; cytosolic; gene therapy; liver disease;
XX proliferative disorder; renal failure; cardiovascular disorder; defect;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
XX OS Homo sapiens.
XX
XX PN WO200257460-A2.
XX
XX PD 25-JUL-2002.
XX
XX PF 20-DEC-2001; 2001WO-US050459.
XX
XX PR 20-DEC-2000; 2000US-0256688P.
XX
XX PR 30-MAR-2001; 2001US-0280186P.
XX
XX PR 01-MAY-2001; 2001US-0287735P.
XX
XX PR 05-JUN-2001; 2001US-0295848P.
XX
XX PR 25-JUN-2001; 2001US-0300465P.
XX
XX PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
XX PI Siewers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
XX PI Kyrsteck S, Mcatee P, Suchard S, Banas D;
XX
XX DR WPI; 2002-599721/64.
XX
XX DR N-PSDB; ACC60521.
XX
XX PT Novel polynucleotides encoding human phosphatase polypeptides useful in
XX the prevention or treatment of e.g. proliferative and cardiovascular
XX disorders.
XX
XX PS Claim 5; Fig 5; 801pp; English.
XX
XX SQ Sequence 665 AA;
XX
XX Query Match 98.7%; Score 1544; DB 5; Length 665;
XX Best Local Similarity 99.0%; Pred. No. 5.7e-162;
XX Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNTSHIILPAININCSKLMKRRLQ 60
XX |||||||
XX DB 1 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNTSHIILPAININCSKLMKRRLQ 60
XX
XX QY 61 DKVLTIELIOHSAKHKVIDCSQKVVVYDOSQDVASLSDFLTVLGKLEKSFNSVHL 120
XX |||||||
XX DB 61 DKVLTIELIOHSAKHKVIDCSQKVVVYDOSQDVASLSDFLTVLGKLEKSFNSVHL 120

```

QY 121 LAGGFAPFRCFPGLCCEKSTLVPCTISQPCLPVANIGPRLIPNLYLGGQDVLNKKEL 180
 DB 121 LAGGFAPFRCFPGLCCEKSTLVPCTISQPCLPVANIGPRLIPNLYLGGQDVLNKKEL 180
 QY 181 QONGIGYVLANASTCPKDPFIPESHFLRPVNDSPCEKILPMDKSDVFTEKKAANGCV 240
 DB 181 QONGIGYVLANASTCPKDPFIPESHFLRPVNDSPCEKILPMDKSDVFTEKKAANGCV 240
 QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPISNPNFLGQLDYEEKIKN 300
 DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPISNPNFLGQLDYEEKIKN 300
 QY 301 QT 302
 DB 301 QT 302

RESULT 15
 AAM25744
 ID AAM25744 standard; protein; 672 AA.
 XX
 AC AAM25744;
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:1259.
 XX
 DE Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KM antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KM antibacterial; endocrine; cardiant; central nervous system; virucide;
 KM anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 KM antitubercular; haemostatic; vulnery; antitumor; osteopathic; eczema;
 KM dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;
 KM neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KM antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KM genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KM chromocytoma; osteoporosis; severe combined immunodeficiency;
 KM allergic rhinitis; diabetes; multiple sclerosis; depression;
 KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KM neurological disorder.
 XX
 OS Homo sapiens.
 XX
 PN W0200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 22-DEC-2000; 2000MO-US035017.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-457603/49.
 DR N-PSDB; AAM99685.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
 XX
 PS Claim 20; Page 260; 1217pp; English.
 XX
 CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and cells
 CC they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;

CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
 CC antitumor; osteopathic; dermatological; antiarthritic; antiaesthetic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic disease, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders
 XX

SQ Sequence 672 AA;

Query Match 98.7%; Score 1544; DB 4; Length 672;
 Best Local Similarity 99.0%; Pred. No. 5,8e-162;
 Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAHEMIGTOYTERLVALLESGETEKLIDSRPFVEYNTSHLEAININCSKMKRRLQ 60
 DB 8 MAHEMIGTOYTERGVALLESGETEKLIDSRPFVEYNTSHLEAININCSKMKRRLQ 67
 QY 61 DKVLITELIQHSAKHVDIDSGQVVVYDQSSQDVASLSDCEFLVLLGKLEKSFNSVHL 120
 DB 68 DKVLITELIQHSAKHVDIDSGQVVVYDQSSQDVASLSDCEFLVLLGKLEKSFNSVHL 127
 QY 121 LAGGFAPFRCFPGLCCEKSTLVPCTISQPCLPVANIGPRLIPNLYLGGQDVLNKKEL 180
 DB 128 LAGGFAPFRCFPGLCCEKSTLVPCTISQPCLPVANIGPRLIPNLYLGGQDVLNKKEL 187
 QY 181 QONGIGYVLANASTCPKDPFIPESHFLRPVNDSPCEKILPMDKSDVFTEKKAANGCV 240
 DB 188 QONGIGYVLANASTCPKDPFIPESHFLRPVNDSPCEKILPMDKSDVFTEKKAANGCV 247
 QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPISNPNFLGQLDYEEKIKN 300
 DB 248 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPISNPNFLGQLDYEEKIKN 307
 QY 301 QT 302
 DB 308 QT 309

Search completed: June 21, 2004, 13:20:33
 Job time : 33.7291 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:19:48 ; Search time 25.9214 Seconds
(without alignments)
3289.124 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMGTGIVTERLVALL.....PNNFGLGDLDEKIKKQF 302

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1163542 segs, 28213646 residues
Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1552	99.2	662 12 US-10-072-012-258	Sequence 258, App
2	1552	99.2	665 9 US-09-816-494-2	Sequence 2, Appl
3	1552	99.2	665 9 US-09-964-277-2	Sequence 2, Appl
4	1552	99.2	665 12 US-10-072-012-680	Sequence 680, App
5	1552	99.2	665 12 US-10-168-506-14	Sequence 14, Appl
6	1552	99.2	665 12 US-10-343-357-7	Sequence 7, Appl
7	1552	99.2	665 15 US-10-377-072-26	Sequence 26, Appl
8	1552	99.2	665 16 US-10-257-026-2	Sequence 2, Appl
9	1552	99.2	665 16 US-10-648-593-240	Sequence 240, App
10	1552	99.2	665 16 US-10-648-593-247	Sequence 247, App
11	1552	99.2	690 12 US-10-072-012-679	Sequence 679, App
12	1552	99.2	690 12 US-10-072-012-703	Sequence 703, App
13	1552	99.2	690 12 US-10-825-114-54204	Sequence 54204, A
14	1545	98.8	665 12 US-10-072-012-681	Sequence 681, App
15	1545	98.8	665 15 US-10-094-749-2312	Sequence 2312, App

16	1544	98.7	672 12 US-10-296-115-1259	Sequence 1259, App
17	1517	97.0	662 12 US-10-072-012-682	Sequence 682, App
18	1514	96.8	677 12 US-10-072-012-683	Sequence 683, App
19	1504.5	96.2	680 12 US-10-072-012-256	Sequence 256, App
20	988	63.2	625 12 US-10-072-012-659	Sequence 659, App
21	985	63.0	663 12 US-10-072-012-700	Sequence 700, App
22	837	53.5	616 12 US-10-072-012-266	Sequence 266, App
23	807	51.6	155 9 US-09-964-277-7	Sequence 7, Appl
24	654.5	41.8	169 14 US-10-346-356-15	Sequence 15, Appl
25	654.5	41.8	170 9 US-09-775-925-26	Sequence 26, Appl
26	654.5	41.8	170 9 US-09-847-519A-11	Sequence 11, Appl
27	654.5	41.8	170 12 US-10-655-073-16	Sequence 16, Appl
28	654.5	41.8	170 14 US-10-314-058-14	Sequence 14, Appl
29	654.5	41.8	170 14 US-10-405-808-16	Sequence 16, Appl
30	646	41.3	517 9 US-09-964-277-21	Sequence 21, Appl
31	631.5	40.4	461 12 US-10-072-012-701	Sequence 701, App
32	620	39.6	501 12 US-10-072-012-702	Sequence 702, App
33	619	39.6	155 9 US-09-964-277-6	Sequence 6, Appl
34	619	39.6	155 9 US-09-955-732-6	Sequence 6, Appl
35	469	30.0	444 9 US-09-964-899-47	Sequence 47, Appl
36	469	30.0	482 12 US-10-058-270A-130	Sequence 130, App
37	469	30.0	482 14 US-10-346-356-2	Sequence 2, Appl
38	469	30.0	482 16 US-10-648-593-164	Sequence 164, App
39	462.5	29.6	381 14 US-10-184-832-2	Sequence 2, Appl
40	426	27.2	394 12 US-10-060-065-25	Sequence 25, Appl
41	426	27.2	394 14 US-10-059-585-46	Sequence 46, Appl
42	425	27.2	394 9 US-09-736-457-805	Sequence 805, App
43	425	27.2	394 9 US-09-736-457-827	Sequence 827, App
44	425	27.2	394 9 US-09-902-941-805	Sequence 805, App
45	425	27.2	394 9 US-09-902-941-827	Sequence 827, App

ALIGNMENTS

RESULT 1
US-10-072-012-258
; Sequence 258, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tcheurev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerbussen, Bryan
; APPLICANT: Batturajan, Meera
; APPLICANT: Shmirek, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Beba
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Futrak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31

```

; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 258
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-258
```

```

Query Match          99.2%; Score 1552; DB 12; Length 662;
Best Local Similarity 99.3%; Pred. No. 3,4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
DB      1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
QY      61 DKVLITELIHSARKKVIDCSQKVYVYDSSQDVASLSSDCEFLTVLLGKLEKSFNSVHL 120
DB      61 DKVLITELIHSARKKVIDCSQKVYVYDSSQDVASLSSDCEFLTVLLGKLEKSFNSVHL 120
QY      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
DB      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
QY      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
DB      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
QY      181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
DB      181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
QY      241 LVHCLAGISRSATTAIAIYIMKRMDSLDEAYRFVKEKRPPTISPNNFLGQLLDYEKKIKN 300
DB      241 LVHCLAGISRSATTAIAIYIMKRMDSLDEAYRFVKEKRPPTISPNNFLGQLLDYEKKIKN 300
QY      301 QT 302
DB      301 QT 302
```

```

RESULT 2
US-09-816-494-2
; Sequence 2, Application us/09816494
; Patent No. US20020034807A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10446-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-494-2
```

```

Query Match          99.2%; Score 1552; DB 9; Length 665;
Best Local Similarity 99.3%; Pred. No. 3,4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
DB      1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
QY      61 DKVLITELIHSARKKVIDCSQKVYVYDSSQDVASLSSDCEFLTVLLGKLEKSFNSVHL 120
DB      61 DKVLITELIHSARKKVIDCSQKVYVYDSSQDVASLSSDCEFLTVLLGKLEKSFNSVHL 120
QY      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
DB      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
QY      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
DB      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
QY      181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
DB      181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
QY      241 LVHCLAGISRSATTAIAIYIMKRMDSLDEAYRFVKEKRPPTISPNNFLGQLLDYEKKIKN 300
DB      241 LVHCLAGISRSATTAIAIYIMKRMDSLDEAYRFVKEKRPPTISPNNFLGQLLDYEKKIKN 300
QY      301 QT 302
DB      301 QT 302
```

```

RESULT 3
US-09-964-277-2
; Sequence 2, Application us/09964277
; Patent No. US20020137170A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125,434
; CURRENT APPLICATION NUMBER: US/09/964,277
; PRIOR FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-964-277-2
```

```

Query Match          99.2%; Score 1552; DB 9; Length 665;
Best Local Similarity 99.3%; Pred. No. 3,4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
DB      1 MAHEMIGTQIVTERVALLESSTGTEKVLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
QY      61 DKVLITELIHSARKKVIDCSQKVYVYDSSQDVASLSSDCEFLTVLLGKLEKSFNSVHL 120
DB      61 DKVLITELIHSARKKVIDCSQKVYVYDSSQDVASLSSDCEFLTVLLGKLEKSFNSVHL 120
QY      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
DB      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
QY      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
DB      121 LAGGFAEFSRCFPGICGKSTLVPCTISQPCLPVANIPTIRILPNLYLGGQRDVANKELI 180
QY      181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
DB      181 QONGIGYVLNASYTCPEKDPFIPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKASNGCV 240
QY      241 LVHCLAGISRSATTAIAIYIMKRMDSLDEAYRFVKEKRPPTISPNNFLGQLLDYEKKIKN 300
DB      241 LVHCLAGISRSATTAIAIYIMKRMDSLDEAYRFVKEKRPPTISPNNFLGQLLDYEKKIKN 300
QY      301 QT 302
```

Db 301 QT 302

RESULT 4

US-10-072-012-680
; Sequence 680, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patwrajan, Meera
; APPLICANT: Shinkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlech, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 680
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-680

Query Match 99.2%; Score 1552; DB 12; Length 665;
Best Local Similarity 99.3%; Pred. No 3, 4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
DB 1 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60

QY 61 DKVLITELIQHSAKHVDIDCSQKVVYDQSSQDVASLSSDCLTVLGLKLEKSFNSVHL 120
DB 61 DKVLITELIQHSAKHVDIDCSQKVVYDQSSQDVASLSSDCLTVLGLKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIGPTRILPNLYLGGQDVLNKELI 180
DB 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIGPTRILPNLYLGGQDVLNKELI 180
QY 181 QONGIGYVNASYTCRKPDPFIPESHFLRPVNDSPCEKILPMWLDKSVDFTEKAKASNGCV 240
DB 181 QONGIGYVNASYTCRKPDPFIPESHFLRPVNDSPCEKILPMWLDKSVDFTEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPDISPBNFLGQLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPDISPBNFLGQLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

RESULT 5

US-10-168-506-14
; Sequence 14, Application US/10168506
; Publication No. US20040053229A1
; GENERAL INFORMATION:
; APPLICANT: PLOMAN, GREGORY D.
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERRARD
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: HILL, RON
; APPLICANT: PLANAGAN, BETTER
; TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
; FILE REFERENCE: 038602/1351
; CURRENT APPLICATION NUMBER: US/10/168,506
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: PCT/US00/34736
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-168-506-14

Query Match 99.2%; Score 1552; DB 12; Length 665;
Best Local Similarity 99.3%; Pred. No 3, 4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
DB 1 MAHEMIGTQIVTERLVALLESGETEKLIDSRPFVEYNTSHILEAININCSKLMKRLQ 60
QY 61 DKVLITELIQHSAKHVDIDCSQKVVYDQSSQDVASLSSDCLTVLGLKLEKSFNSVHL 120
DB 61 DKVLITELIQHSAKHVDIDCSQKVVYDQSSQDVASLSSDCLTVLGLKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIGPTRILPNLYLGGQDVLNKELI 180
DB 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQPCLPVANIGPTRILPNLYLGGQDVLNKELI 180
QY 181 QONGIGYVNASYTCRKPDPFIPESHFLRPVNDSPCEKILPMWLDKSVDFTEKAKASNGCV 240
DB 181 QONGIGYVNASYTCRKPDPFIPESHFLRPVNDSPCEKILPMWLDKSVDFTEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPDISPBNFLGQLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPDISPBNFLGQLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302

Db 301 QT 302

RESULT 6

US-10-343-357-7
; Sequence 7, Application US/10343357
; Publication No. US20040058341A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom
; APPLICANT: ELIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
; APPLICANT: LEE, Ernestine A.; HAFALIA, April J. A.
; APPLICANT: LU, Dying Aina M.; TRIBOLET, Catherine M.
; APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
; APPLICANT: YUE, Henry; WARREN, Bridget A.
; APPLICANT: NGUYEN, Daniel B.; CHAMLA, Nandinder K.
; APPLICANT: KEARNEY, Liam
; TITLE OF INVENTION: PROTEIN PHOSPHATASES
; FILE REFERENCE: PI-0173 PCT
; CURRENT APPLICATION NUMBER: US/10/343,357
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: PCT/US01/23716
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,679
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/223,272
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/224,309
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 60/226,728
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/229,254
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: US 60/231,366
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CD1
US-10-343-357-7

Query Match 99.2%; Score 1552; DB 12; Length 665;
Best Local Similarity 99.3%; Pred. No. 3.4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 MAHEMIGTOYTERVALLESGETEKVLLIDSRPFVYNTSHILEAININCSKLMRRLOQ 60
1 MAHEMIGTOYTERVALLESGETEKVLLIDSRPFVYNTSHILEAININCSKLMRRLOQ 60
61 DKVLTTELIOHSAKHKVIDCSQKVVVYDQSSQVVASLSSDCEFLVLLGKLEKSFNSVHL 120
61 DKVLTTELIOHSAKHKVIDCSQKVVVYDQSSQVVASLSSDCEFLVLLGKLEKSFNSVHL 120
121 LAGFAEFSCFCPELCEGKSTLVPTCISQCLPVANIGPTRILPMLYGCQRDVNLKELI 180
121 LAGFAEFSCFCPELCEGKSTLVPTCISQCLPVANIGPTRILPMLYGCQRDVNLKELI 180
121 LAGFAEFSCFCPELCEGKSTLVPTCISQCLPVANIGPTRILPMLYGCQRDVNLKELI 180
181 QONGIGYVLNASTCPKPDFIPESHFLRPVNDSPCEKILPWLKSDVDFIEKAKANGCV 240
181 QONGIGYVLNASTCPKPDFIPESHFLRPVNDSPCEKILPWLKSDVDFIEKAKANGCV 240
241 LVHCLAGISSATTAIAYIMKRDMSLDEAYRFVKEKRPITISPNFNLGQLLDYEKIKN 300
241 LVHCLAGISSATTAIAYIMKRDMSLDEAYRFVKEKRPITISPNFNLGQLLDYEKIKN 300
301 QT 302

Db 301 QT 302

RESULT 7

US-10-377-072-26
; Sequence 26, Application US/10377072
; Publication No. US20040009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Miyoung
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; FILE REFERENCE: MPI03-0180NMIM
; CURRENT APPLICATION NUMBER: US/10/377,072
; PRIOR FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-377-072-26

Query Match 99.2%; Score 1552; DB 15; Length 665;
Best Local Similarity 99.3%; Pred. No. 3.4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 MAHEMIGTOYTERVALLESGETEKVLLIDSRPFVYNTSHILEAININCSKLMRRLOQ 60
1 MAHEMIGTOYTERVALLESGETEKVLLIDSRPFVYNTSHILEAININCSKLMRRLOQ 60
61 DKVLTTELIOHSAKHKVIDCSQKVVVYDQSSQVVASLSSDCEFLVLLGKLEKSFNSVHL 120
61 DKVLTTELIOHSAKHKVIDCSQKVVVYDQSSQVVASLSSDCEFLVLLGKLEKSFNSVHL 120
121 LAGFAEFSCFCPELCEGKSTLVPTCISQCLPVANIGPTRILPMLYGCQRDVNLKELI 180
121 LAGFAEFSCFCPELCEGKSTLVPTCISQCLPVANIGPTRILPMLYGCQRDVNLKELI 180
121 LAGFAEFSCFCPELCEGKSTLVPTCISQCLPVANIGPTRILPMLYGCQRDVNLKELI 180
181 QONGIGYVLNASTCPKPDFIPESHFLRPVNDSPCEKILPWLKSDVDFIEKAKANGCV 240
181 QONGIGYVLNASTCPKPDFIPESHFLRPVNDSPCEKILPWLKSDVDFIEKAKANGCV 240
Db 181 QONGIGYVLNASTCPKPDFIPESHFLRPVNDSPCEKILPWLKSDVDFIEKAKANGCV 240


```

Oy 241 LVHCLAGISRSATTAIAIYIMGRMMSLDEAYRFYKRRPTISPNFNLGQLDYEKKIKN 300
Db 241 LVHCLAGISRSATTAIAIYIMGRMMSLDEAYRFYKRRPTISPNFNLGQLDYEKKIKN 300
Oy 301 QT 302
Db 301 QT 302

RESULT 8
US-10-257-026-2
; Sequence 2, Application US/10257026
; Publication No. US20040086859A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10KDM5
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-257-026-2

Query Match 99.2%; Score 1552; DB 16; Length 665;
Best Local Similarity 99.3%; Pred. No. 3,4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MAHEMIGTOYITERLVALLESGETEKVLLIDSRPFVEYNTSHILDAININCSKIMKRLQ 60
Db 1 MAHEMIGTOYITERLVALLESGETEKVLLIDSRPFVEYNTSHILDAININCSKIMKRLQ 60
Oy 61 DKVLITELIOHSARKKVDIDCSOKVWVYDSSODVASISDPCFLTLLGKLEKSPNSVHL 120
Db 61 DKVLITELIOHSARKKVDIDCSQKVVYDSSQDVASISDPCFLTLLGKLEKSPNSVHL 120
Oy 121 LAGGFAERSRCFPGICEGSKTLVPTCISQPCLPVANIPTIRILPNLYLGCORVLANKELI 180
Db 121 LAGGFAERSRCFPGICEGSKTLVPTCISQPCLPVANIPTIRILPNLYLGCORVLANKEIM 180
Oy 181 QONGIGYVLANSYTCPKDPFIPESHFLRVPNVDSFCEKILFMDLKSVDPIEKAYASNGCV 240
Db 181 QONGIGYVLANSNCTCPKDPFIPESHFLRVPNVDSFCEKILFMDLKSVDPIEKAYASNGCV 240
Oy 241 LVHCLAGISRSATTAIAIYIMGRMMSLDEAYRFYKRRPTISPNFNLGQLDYEKKIKN 300
Db 241 LVHCLAGISRSATTAIAIYIMGRMMSLDEAYRFYKRRPTISPNFNLGQLDYEKKIKN 300
Oy 301 QT 302
Db 301 QT 302

RESULT 9
US-10-648-593-240
; Sequence 240, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 240

```

```

: LENGTH: 665;
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-648-593-240

Query Match          99.2%; Score 1552; DB 16; Length 665;
Best Local Similarity 99.3%; Pred. No. 3,4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
DB      1 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60

QY      61 DKVLITELLIGHSAHKKVDIDCSQKVVYDDSSQDVASLSDCFLTVLLGLGLESFNSVHL 120
DB      61 DKVLITELLIGHSAHKKVDIDCSQKVVYDDSSQDVASLSDCFLTVLLGLGLESFNSVHL 120

QY      121 LAGFAFAFSRCFPGLCGKSTLVPTCISQCLFVNANIGPRILIPNLVYLGCQDVLNKEML 180
DB      121 LAGFAFAFSRCFPGLCGKSTLVPTCISQCLFVNANIGPRILIPNLVYLGCQDVLNKEML 180

QY      181 QONGIGVYLAASATYCPKPDFIPESHFLRVVNDSPCEKILPWLDSVDFTEKAKASNGCV 240
DB      181 QONGIGVYLAASATYCPKPDFIPESHFLRVVNDSPCEKILPWLDSVDFTEKAKASNGCV 240

QY      241 LVHCLAGISSATTAIAYIMKRWMDSLDEAYRVRVKEKRPITSNENFLGOLLVEKKIKY 300
DB      241 LVHCLAGISSATTAIAYIMKRWMDSLDEAYRVRVKEKRPITSNENFLGOLLVEKKIKY 300

QY      301 QT 302
DB      301 QT 302

RESULT 10
US-10-648-593-247
: Sequence 247, Application US/10648593
: Publication No. US20040106132A1
: GENERAL INFORMATION:
: APPLICANT: Bristol-Myers Squibb Company
: TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THATM
: TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
: TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
: FILE REFERENCE: D0273 NP
: CURRENT APPLICATION NUMBER: US/10/648,593
: CURRENT FILING DATE: 2003-08-26
: PRIOR APPLICATION NUMBER: 60/406,385
: PRIOR FILING DATE: 2002-08-27
: NUMBER OF SEQ ID NOS: 557
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 247
: LENGTH: 665
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-648-593-247

Query Match          99.2%; Score 1552; DB 16; Length 665;
Best Local Similarity 99.3%; Pred. No. 3,4e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60
DB      1 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLQ 60

QY      61 DKVLITELLIGHSAHKKVDIDCSQKVVYDDSSQDVASLSDCFLTVLLGLGLESFNSVHL 120
DB      61 DKVLITELLIGHSAHKKVDIDCSQKVVYDDSSQDVASLSDCFLTVLLGLGLESFNSVHL 120

QY      121 LAGFAFAFSRCFPGLCGKSTLVPTCISQCLFVNANIGPRILIPNLVYLGCQDVLNKEML 180
DB      121 LAGFAFAFSRCFPGLCGKSTLVPTCISQCLFVNANIGPRILIPNLVYLGCQDVLNKEML 180

QY      181 QONGIGVYLAASATYCPKPDFIPESHFLRVVNDSPCEKILPWLDSVDFTEKAKASNGCV 240
DB      181 QONGIGVYLAASATYCPKPDFIPESHFLRVVNDSPCEKILPWLDSVDFTEKAKASNGCV 240

```

```

Db      181  QONGGYVLMASNTCPKDPFIPESHFLRPVNDSECEKILPWLDKSVDFIEKAKASNGCV 240
Qy      241  LVHCLAGISRSATTAIAYIMKRMDSLDEAYRFVKEKRPPTISPNFNLGQLDYEKKIKN 300
Db      241  LVHCLAGISRSATTAIAYIMKRMDSLDEAYRFVKEKRPPTISPNFNLGQLDYEKKIKN 300
Qy      301  QT 302
Db      301  QT 302

```

```

RESULT 11
US-10-072-012-679
/ Sequence 679, Application US/10072012
/ Publication No. US20040033493A1
/ GENERAL INFORMATION:
/ APPLICANT: Tcherenev, Velizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Patuturajan, Meera
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangolli, Bsha
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Colman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grose, William M.
/ APPLICANT: Alsobrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072.012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 60/266,767
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: 60/267,057
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/266,975
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/267,459
/ PRIOR FILING DATE: 2001-02-08
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1391
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 679
/ LENGTH: 690
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-072-012-679

```

```

Query Match      99.2%; Score 1552; DB 12; Length 690;
Best Local Similarity 99.3%; Pred. No. 3.6e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1  MAHMMIGTQVTERLVALLSSGTBKYLLIDSRFPVENYNTSHILEAININISKLMKRLQ 60
Db      26  MAHMMIGTQVTERLVALLSSGTBKYLLIDSRFPVENYNTSHILEAININISKLMKRLQ 85
Qy      61  DKVLITELIHSAAKHVVDICSOQVWVYDOSODVASLSSDCPLVTLGLKESFNSVHL 120
Db      86  DKVLITELIHSAAKHVVDICSOQVWVYDOSODVASLSSDCPLVTLGLKESFNSVHL 145
Qy      121  LAGFAEFSSCFPGLCGKSTLVPTCISQCLPVANIGPTRLIPNLVYLGQQRDVNLKELI 180
Db      146  LAGFAEFSSCFPGLCGKSTLVPTCISQCLPVANIGPTRLIPNLVYLGQQRDVNLKELI 205
Qy      181  QONGIGVLMASNTCPKDPFIPESHFLRPVNDSECEKILPWLDKSVDFIEKAKASNGCV 240
Db      206  QONGIGVLMASNTCPKDPFIPESHFLRPVNDSECEKILPWLDKSVDFIEKAKASNGCV 265
Qy      241  LVHCLAGISRSATTAIAYIMKRMDSLDEAYRFVKEKRPPTISPNFNLGQLDYEKKIKN 300
Db      266  LVHCLAGISRSATTAIAYIMKRMDSLDEAYRFVKEKRPPTISPNFNLGQLDYEKKIKN 325
Qy      301  QT 302
Db      326  QT 327

```

```

RESULT 12
US-10-072-012-703
/ Sequence 703, Application US/10072012
/ Publication No. US20040033493A1
/ GENERAL INFORMATION:
/ APPLICANT: Tcherenev, Velizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Patuturajan, Meera
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangolli, Bsha
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Colman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grose, William M.
/ APPLICANT: Alsobrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072.012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02

```

PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PAM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 703
LENGTH: 690
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-703

Query Match 99.2%; Score 1552; DB 12; Length 690;
Best Local Similarity 99.3%; Pred. No. 3.6e-147;
Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNISHIIEAININCSKLMKRLQ 60
DB 26 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNISHIIEAININCSKLMKRLQ 85
QY 61 DKVLITELIOMSAKHKVDIDCSQKVVYDOSQDVASLSSDCLFVLLGLKESFNSVHL 120
DB 86 DKVLITELIOMSAKHKVDIDCSQKVVYDOSQDVASLSSDCLFVLLGLKESFNSVHL 145
QY 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQCLPVPANIGPTRLIPNLVLLGCGQDVLNKELI 180
DB 146 LAGGFAEFSRCFPGCEGKSTLVPTCISQCLPVPANIGPTRLIPNLVLLGCGQDVLNKELI 205
QY 181 QONGIGYVNLASTYTCRPPDIPESHFLRVVNDSPCEKILPMDKSVDFEKAASNGCV 240
DB 206 QONGIGYVNLASTYTCRPPDIPESHFLRVVNDSPCEKILPMDKSVDFEKAASNGCV 265
QY 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFPLGQLDYEKKIKN 300
DB 266 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFPLGQLDYEKKIKN 325
QY 301 QT 302
DB 326 QT 327

RESULT 13

US-10-425-114-54204
Sequence 54204, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaka, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 54204
LENGTH: 690
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Clone ID: LIB4119-028-H6_FLI.pep
US-10-425-114-54204
Query Match 99.2%; Score 1552; DB 12; Length 690;
Best Local Similarity 99.3%; Pred. No. 3.6e-147;

Matches 300; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNISHIIEAININCSKLMKRLQ 60
DB 26 MAHEMIGTQIVTERLVALLSGETEKVLLIDSRPFVEYNISHIIEAININCSKLMKRLQ 85
QY 61 DKVLITELIOMSAKHKVDIDCSQKVVYDOSQDVASLSSDCLFVLLGLKESFNSVHL 120
DB 86 DKVLITELIOMSAKHKVDIDCSQKVVYDOSQDVASLSSDCLFVLLGLKESFNSVHL 145
QY 121 LAGGFAEFSRCFPGCEGKSTLVPTCISQCLPVPANIGPTRLIPNLVLLGCGQDVLNKELI 180
DB 146 LAGGFAEFSRCFPGCEGKSTLVPTCISQCLPVPANIGPTRLIPNLVLLGCGQDVLNKELI 205
QY 181 QONGIGYVNLASTYTCRPPDIPESHFLRVVNDSPCEKILPMDKSVDFEKAASNGCV 240
DB 206 QONGIGYVNLASTYTCRPPDIPESHFLRVVNDSPCEKILPMDKSVDFEKAASNGCV 265
QY 241 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFPLGQLDYEKKIKN 300
DB 266 LVHCLAGISRSATITAIAYIMKMDMSLDEAYRFVKEKRPITISPNFPLGQLDYEKKIKN 325
QY 301 QT 302
DB 326 QT 327

RESULT 14

US-10-072-012-681
Sequence 681, Application US/10072012
Publication No. US2004003493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Patuturajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Beha
APPLICANT: Radigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie E.
APPLICANT: Taupier, Jr., Raymond J.
APPLICANT: Gusev, Vladimiro Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine B.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05

```
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 681
; LENGTH: 665
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-072-012-681
```

```
Query Match          98.8%; Score 1545; DB 12; Length 665;
Best Local Similarity 99.0%; Pred. No. 1.7e-146;
Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 MAHEMIGTQVTERLVALLESGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLESGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIQHSARKHVIDIDCSQKVVYDQSDVASLSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIQHSARKHVIDIDCSQKVVYDQSDVASLSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCCEKSTLVPTCISQPCLPVANIGPTRILPMLYLGCGRDVNLKELI 180
DB 121 LAGGFAEFSRCFPGLCCEKSTLVPTCISQPCLPVANIGPTRILPMLYLGCGRDVNLKELI 180
QY 181 QONGIGVNLASNTCPKPDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGVNLASNTCPKPDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
```

```
RESULT 15
US-10-094-749-2312
; Sequence 2312, Application US/10094749
; Publication No. US20030219741A1
```

GENERAL INFORMATION:

```
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YUKI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHITO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
```

```
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2312
; LENGTH: 665
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-094-749-2312
```

```
Query Match          98.8%; Score 1545; DB 15; Length 665;
Best Local Similarity 99.0%; Pred. No. 1.7e-146;
Matches 299; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 MAHEMIGTQVTERLVALLESGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
DB 1 MAHEMIGTQVTERLVALLESGETEKVLLIDSRPFVEYNTSHILEAININCSKLMKRRLOQ 60
QY 61 DKVLITELIQHSARKHVIDIDCSQKVVYDQSDVASLSDCFLTVLLGKLEKSFNSVHL 120
DB 61 DKVLITELIQHSARKHVIDIDCSQKVVYDQSDVASLSDCFLTVLLGKLEKSFNSVHL 120
QY 121 LAGGFAEFSRCFPGLCCEKSTLVPTCISQPCLPVANIGPTRILPMLYLGCGRDVNLKELI 180
DB 121 LAGGFAEFSRCFPGLCCEKSTLVPTCISQPCLPVANIGPTRILPMLYLGCGRDVNLKELI 180
QY 181 QONGIGVNLASNTCPKPDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
DB 181 QONGIGVNLASNTCPKPDFIPESHFLRPVNDSCFCEKILPMLDKSVDFIEKAKASNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIKN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIKN 300
QY 301 QT 302
DB 301 QT 302
```

```
Search completed: June 21, 2004, 13:24:58
Job time : 26.9214 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:20:38 ; Search time 10.3061 Seconds
(without alignments)
1512.797 Million cell updates/sec

```
Title: US-10-029-345A-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMIGTQIVTERLVALLE.....PWFNFGQLLDYKKIKNOT 3022
```

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Searched: 389414 begs, 51625971 residues
Total number of hits satisfying chosen parameters: 389414

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
IssuePatents AA:*
1: /cn2_6/prodata/2/iaa/5A_COMB.pcp.*
2: /cn2_6/prodata/2/iaa/5B_COMB.pcp.*
3: /cn2_6/prodata/2/iaa/6A_COMB.pcp.*
4: /cn2_6/prodata/2/iaa/6B_COMB.pcp.*
5: /cn2_6/prodata/2/iaa/POTUS_COMB.pcp.*
6: /cn2_6/prodata/2/iaa/backfill1.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1552	99.2	665	4	US-09-816-994-2	Sequence 2, Appl1
2	654.5	41.8	170	4	US-09-544-716-14	Sequence 14, Appl1
3	654.5	41.8	170	4	US-09-557-921-15	Sequence 15, Appl1
4	654.5	41.8	170	4	US-09-564-357-17	Sequence 17, Appl1
5	654.5	41.8	170	4	US-09-619-380-16	Sequence 16, Appl1
6	469	30.0	482	4	US-09-557-921-2	Sequence 2, Appl1
7	425	27.2	314	4	US-09-311-671B-11	Sequence 11, Appl1
8	425	27.2	367	2	US-08-990-379-6	Sequence 6, Appl1
9	425	27.2	394	2	US-08-530-290-23	Sequence 23, Appl1
10	425	27.2	394	4	US-09-702-705-805	Sequence 805, Appl1
11	425	27.2	394	4	US-09-702-705-827	Sequence 827, Appl1
12	425	27.2	394	4	US-09-736-457-805	Sequence 805, Appl1
13	425	27.2	394	4	US-09-736-457-827	Sequence 827, Appl1
14	425	27.2	394	4	US-09-614-124B-805	Sequence 805, Appl1
15	425	27.2	394	4	US-09-614-124B-827	Sequence 827, Appl1
16	425	27.2	394	4	US-09-611-325-805	Sequence 805, Appl1
17	425	27.2	394	4	US-09-611-325-827	Sequence 827, Appl1
18	425	27.2	394	4	US-09-569-184-805	Sequence 805, Appl1
19	425	27.2	394	4	US-09-569-184-827	Sequence 827, Appl1
20	423.5	27.1	395	2	US-08-990-379-5	Sequence 5, Appl1
21	415	26.5	367	2	US-08-530-290-24	Sequence 24, Appl1
22	409.5	26.2	313	2	US-08-990-379-7	Sequence 7, Appl1
23	409	26.2	314	3	US-09-164-193-22	Sequence 22, Appl1
24	409	26.2	314	4	US-09-221-448A-22	Sequence 22, Appl1
25	377.5	24.1	393	2	US-08-990-379-4	Sequence 4, Appl1
26	357	22.8	302	4	US-09-702-705-806	Sequence 806, Appl1
27	357	22.8	302	4	US-09-736-457-806	Sequence 806, Appl1

28	357	22.8	302	4	US-09-614-134B-806	Sequence 806, App
29	357	22.8	302	4	US-09-671-332S-006	Sequence 806, App
30	357	22.8	302	4	US-09-589-184-806	Sequence 806, App
31	349.5	22.3	397	2	US-08-999-379-8	Sequence 8, App1
32	346.5	22.2	168	4	US-09-544-716-13	Sequence 13, App1
33	346.5	22.2	168	4	US-09-557-921-13	Sequence 13, App1
34	346.5	22.2	168	4	US-09-556-357-16	Sequence 16, App1
35	346.5	22.6	168	4	US-09-619-380-15	Sequence 16, App1
36	338.5	21.6	170	4	US-09-544-716-12	Sequence 12, App1
37	338.5	21.6	170	4	US-09-557-921-12	Sequence 12, App1
38	338.5	21.6	170	4	US-09-556-357-15	Sequence 15, App1
39	338.5	21.6	170	4	US-09-619-380-14	Sequence 14, App1
40	319	20.4	169	4	US-09-544-716-16	Sequence 16, App1
41	319	20.4	169	4	US-09-557-921-17	Sequence 17, App1
42	319	20.4	169	4	US-09-556-357-19	Sequence 19, App1
43	319	20.4	169	4	US-09-619-380-18	Sequence 18, App1
44	312	19.9	168	4	US-09-544-716-15	Sequence 15, App1
45	312	19.9	168	4	US-09-557-921-16	Sequence 16, App1

ALIGNMENTS

```

RESULT 1
US-09-816-494-2
Sequence 2, Application US/09816494
Patent No. 6664089
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OF INVENTION: 38652 AND 21117, NOVEL DUAL SPECIFICITY
FILE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: 191,858
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-09-816-494-2.

```

Query Match	Similarity	99.3%	Score 1552;	DB 4;	Length 665;
Best Local	Matches	300;	Conservative	1;	Mismatches 1;
					Indels 0;
					Gaps 0
Qy	1	MAHEMIGTOIVTERVALLBESGTEKVLIDSRPVEYNTSHILEANININGSKLMRRDQ	60		
Db	1	MAHEMIGTOIVTERVALLBESGTEKVLIDSRPVEYNTSHILEANININGSKLMRRDQ	60		
Qy	61	DKVLITELIGHSANGKVDIDCSQKVVYDDOSSOVAISSSCFLTVLLGKLEKSNVHL	120		
Db	61	DKVLITELIGHSANGKVDIDCSQKVVYDDOSSOVAISSSCFLTVLLGKLEKSNVHL	120		
Qy	121	LAGEFAEFSSRCPGLCEGKSTLVPTCTISQPLPVANIGPTRILPLVLYGCORDVANKELI	180		
Db	121	LAGEFAEFSSRCPGLCEGKSTLVPTCTISQPLPVANIGPTRILPLVLYGCORDVANKELI	180		
Qy	181	QONGIGYVLNASYTCPKRDFIPESHFLRPVNDSPCECEKILPWLKSDVFEKAYASNGCV	240		
Db	181	QONGIGYVLNASYTCPKRDFIPESHFLRPVNDSPCECEKILPWLKSDVFEKAYASNGCV	240		
Qy	241	LVHCLAGISGATIAIAYIMKRMMSIDEARFPKEKRPPTSIPNPNFGOLLDVEKKIKN	300		
Db	241	LVHCLAGISGATIAIAYIMKRMMSIDEARFPKEKRPPTSIPNPNFGOLLDVEKKIKN	300		
Qy	301	QT 302			
Db	301	QT 302			

RESULT 2
US-09-544-716-14
; Sequence 14, Application US/09544716
; Patent No. 6492157
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; TITLE OF INVENTION: DSP-9 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.415
; CURRENT APPLICATION NUMBER: US/09/544.716
; CURRENT FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-544-716-14

Query Match 41.8%; Score 654.5; DB 4; Length 170;
Best Local Similarity 72.5%; Pred. No. 1e-63;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;
QY 134 GLCGK-STLVPTCISQPCLPVANIGPTRLPMUYIGQCRDVLNKLIOQNGIGVYLNAS 192
DB 1 GLCGKPAALLPMSLSQPCLPVPSVGLTRILPHLYIGSQDVLNKLMTQNGISYVLNAS 60
QY 193 YTCRKPDPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSA 252
DB 61 NSCPKDPFICESRMRVPINDNCEKLPMLDKSIEFIDAKLSSQCVLVHCLAGISRSA 120
QY 253 TIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIK 299
DB 121 TIAIAYIMKTMGMSDDAYRFVQDRRPSISPNNFLGQLLEYERTLK 167

RESULT 3
US-09-557-921-15
; Sequence 15, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.416
; CURRENT APPLICATION NUMBER: US/09/557.921
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-921-15

Query Match 41.8%; Score 654.5; DB 4; Length 170;
Best Local Similarity 72.5%; Pred. No. 1e-63;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;
QY 134 GLCGK-STLVPTCISQPCLPVANIGPTRLPMUYIGQCRDVLNKLIOQNGIGVYLNAS 192
DB 1 GLCGKPAALLPMSLSQPCLPVPSVGLTRILPHLYIGSQDVLNKLMTQNGISYVLNAS 60
QY 193 YTCRKPDPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSA 252
DB 61 NSCPKDPFICESRMRVPINDNCEKLPMLDKSIEFIDAKLSSQCVLVHCLAGISRSA 120
QY 253 TIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIK 299
DB 121 TIAIAYIMKTMGMSDDAYRFVQDRRPSISPNNFLGQLLEYERTLK 167

RESULT 4

US-09-564-357-17
; Sequence 17, Application US/09564357
; Patent No. 6645753
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-5 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.413
; CURRENT APPLICATION NUMBER: US/09/564.357
; CURRENT FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-564-357-17

Query Match 41.8%; Score 654.5; DB 4; Length 170;
Best Local Similarity 72.5%; Pred. No. 1e-63;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;
QY 134 GLCGK-STLVPTCISQPCLPVANIGPTRLPMUYIGQCRDVLNKLIOQNGIGVYLNAS 192
DB 1 GLCGKPAALLPMSLSQPCLPVPSVGLTRILPHLYIGSQDVLNKLMTQNGISYVLNAS 60
QY 193 YTCRKPDPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSA 252
DB 61 NSCPKDPFICESRMRVPINDNCEKLPMLDKSIEFIDAKLSSQCVLVHCLAGISRSA 120
QY 253 TIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIK 299
DB 121 TIAIAYIMKTMGMSDDAYRFVQDRRPSISPNNFLGQLLEYERTLK 167

RESULT 5
US-09-619-380-16
; Sequence 16, Application US/09619380
; Patent No. 6649391
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-11 DUAL SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.418
; CURRENT APPLICATION NUMBER: US/09/619.380
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-619-380-16

Query Match 41.8%; Score 654.5; DB 4; Length 170;
Best Local Similarity 72.5%; Pred. No. 1e-63;
Matches 121; Conservative 27; Mismatches 18; Indels 1; Gaps 1;
QY 134 GLCGK-STLVPTCISQPCLPVANIGPTRLPMUYIGQCRDVLNKLIOQNGIGVYLNAS 192
DB 1 GLCGKPAALLPMSLSQPCLPVPSVGLTRILPHLYIGSQDVLNKLMTQNGISYVLNAS 60
QY 193 YTCRKPDPFIPESHFLRVVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRSA 252
DB 61 NSCPKDPFICESRMRVPINDNCEKLPMLDKSIEFIDAKLSSQCVLVHCLAGISRSA 120
QY 253 TIAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIK 299
DB 121 TIAIAYIMKTMGMSDDAYRFVQDRRPSISPNNFLGQLLEYERTLK 167

RESULT 6
US-09-557-921-2

```
Sequence 2, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125, 416
; CURRENT APPLICATION NUMBER: US/09/557, 921
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-921-2

Query Match
Best local Similarity 35.9%; Pred. No. 1e-42;
Matches 107; Conservative 65; Mismatches 96; Indels 30; Gaps 7;

QY 27 LLLIDSRPVEYNTSHILEAININCS-KLMKRLQODKVLITELIQ-HSANGKVDIDCSQK 84
Db 173 VIIDCRPFMEYNKSHIQAVHINCADKISRRLQOGKITVLIDISCSEKDSFKRIFSK 232
QY 85 VVVYDSSQDVASLSDFITVLGLKLEKSPNSVHLLAGFAFERSCFGLCEGK----- 139
Db 233 IIVYDENTNPSRWSQPHATIVESLKRGEKPELVKGLSFKQNHENLCONSLOQE 292
QY 140 -----STLVPTCISQPCLP-VANIGPTRIPLNLYGQORDVNLKELIQONGIGY 187
Db 293 CREVGGAASAASLLPPT--PTTPDIENMLPILPFLILGNBQADQDIDTQRLNIGY 350
QY 188 VNAS-----YTCPKPDFIESHFRLRVVNDSPCEKILPWLDSVDFEKAASNGCVLV 242
Db 351 VINVTHTLPYHKEGLP---NYKRLPATDSMNQNRQYFEBAFEIFEBAHQCGKGLLI 406
QY 243 HCLAGSRKATIAIYIMKMDMSLDAYRFVEKEKRTISPNPFLGQLLDYEKKIKN 300
Db 407 HCAQVSRKATIVAIYIMKMTMTDAYRFVKGKRIISPNLFWGQLLEFEDLNN 464

RESULT 7
US-09-371-671B-11
; Sequence 11, Application US/09371671B
; Patent No. 6548743
; GENERAL INFORMATION:
; APPLICANT: Sheen, Jen
; APPLICANT: Chiu, Man-ling
; TITLE OF INVENTION: DUAL-SPECIFICITY MARK PHOSPHATASE AND USES THEREOF
; FILE REFERENCE: 00786/370002
; CURRENT APPLICATION NUMBER: US/09/371, 671B
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: 60/155, 934
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: 60/095, 938
; PRIOR FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-371-671B-11

Query Match
Best local Similarity 34.3%; Pred. No. 3.7e-38;
Matches 104; Conservative 48; Mismatches 129; Indels 22; Gaps 5;

QY 15 LVALLESGETEKVLLIDSRPVEYNTSHILEAININCSKLMKRLQODKVLITELIQHS 74
Db 13 LRALLREGAAGQCLLDRCRFFAFNAGHIAASVNVRFSTIVRRRAKMGLEHIVPNAELR 72
```

```
QY 75 HKYDIDCSQKVVYDSSQDVASLSDFITVLGLKLEKSPNSVHLLAGFAFERSRCF 132
Db 73 GRLLAGATYAVVLLDRSRASLDGAKRDGTLLAAGLCREARSTQVFFLOGGFEASAS 132
QY 133 PGLCEKSKTLVPTCISQPCLPVANI-----GPTRIPLNLYGQORDVNL 175
Db 133 PELCSKQST--PTGLSLPLSTSVDBASGSCSCSTPLVDGQGFVEILSFYLGSAVHAS 190
QY 176 NKELIQONGIGYVLNASYTCPKPDFIESHFRLRVVNDSPCEKILPWLDSVDFIEKAKA 235
Db 191 RKMDLALGITLALINVSANCPN-HFEGHYQYKSIPEVDNHRKADISSWFNEALDIFDISIKD 249
QY 236 SNGCVLVHCLAGSRKATIAIYIMKMDMSLDAYRFVEKEKRTISPNPFLGQLLDYE 295
Db 250 AGGRVVFHCOAGISRSATTCATLMTKTNVKLDEAFEFVKORRSIISPNFSFGQLLOFE 309
QY 296 KKI 298
Db 310 SQV 312

RESULT 8
US-08-990-379-6
; Sequence 6, Application US/08990379
; Patent No. 5998188
; GENERAL INFORMATION:
; APPLICANT: Stork, Phillip J
; APPLICANT: Maira-Press, Anita
; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
; FILE REFERENCE: 4104-000322USA
; CURRENT APPLICATION NUMBER: US/08/990, 379
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: PCT/US96/10402
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000, 263
; EARLIER FILING DATE: 1995-06-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Mus sp.
US-08-990-379-6

Query Match
Best local Similarity 34.3%; Pred. No. 4.6e-38;
Matches 104; Conservative 48; Mismatches 129; Indels 22; Gaps 5;

QY 15 LVALLESGETEKVLLIDSRPVEYNTSHILEAININCSKLMKRLQODKVLITELIQHS 74
Db 13 LRALLREGAAGQCLLDRCRFFAFNAGHIAASVNVRFSTIVRRRAKMGLEHIVPNAELR 72
QY 75 HKYDIDCSQKVVYDSSQDVASLSDFITVLGLKLEKSPNSVHLLAGFAFERSRCF 132
Db 73 GRLLAGATYAVVLLDRSRASLDGAKRDGTLLAAGLCREARSTQVFFLOGGFEASAS 132
QY 133 PGLCEKSKTLVPTCISQPCLPVANI-----GPTRIPLNLYGQORDVNL 175
Db 133 PELCSKQST--PTGLSLPLSTSVDBASGSCSCSTPLVDGQGFVEILSFYLGSAVHAS 190
QY 176 NKELIQONGIGYVLNASYTCPKPDFIESHFRLRVVNDSPCEKILPWLDSVDFIEKAKA 235
Db 191 RKMDLALGITLALINVSANCPN-HFEGHYQYKSIPEVDNHRKADISSWFNEALDIFDISIKD 249
QY 236 SNGCVLVHCLAGSRKATIAIYIMKMDMSLDAYRFVEKEKRTISPNPFLGQLLDYE 295
Db 250 AGGRVVFHCOAGISRSATTCATLMTKTNVKLDEAFEFVKORRSIISPNFSFGQLLOFE 309
QY 296 KKI 298
Db 310 SQV 312
```


RESULT 9

US-08-530-290-23
Sequence 23, Application US/08530290
Patent No. 5958721
GENERAL INFORMATION:
APPLICANT: Marshall, Christopher John
APPLICANT: Ashworth, Alan
APPLICANT: Hughes, David Anthony
TITLE OF INVENTION: Methods for Screening of Substances for
NUMBER OF SEQUENCES: 24
TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,290
FILING DATE: 14-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB94/00694
FILING DATE: 31-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9402573.1
FILING DATE: 10-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9307250.2
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 084611-000000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-530-290-23

Query Match 27.2%; Score 425; DB 2; Length 394;
Best Local Similarity 33.1%; Pred. No. 5.2e-38;

Matches 99; Conservative 62; Mismatches 116; Indels 22; Gaps 8;

QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCSKLMKRLOQDVLITELI--QHSAXKH 76
DB 39 LPSSG-GKCLLDRCPLLAHSAGYILGSVNRCTIVRR-AKGSVSLQILPABEEVRAR 96
QY 77 VDIDCSQKVVVYDOSQDVASLSDCFLTVLGLKLEKSF--NSVHLIAGFAEFRCFPG 134
DB 97 LRSGLYAVIVYDERSPRASLRSDSTVSLVQALRRNARERTDILCKGGERFSSEYPE 156
QY 135 LCEGKSTL-----VPTCISOP-----CLPVA-----NIGPTRILPNLYGCGQDVLINKEL 179
DB 157 FCSKTKALAIIPPPVPASATPELDLDSCSGCTPLHDQEGVEILPFLYIGSAVHAARRDM 216
QY 180 IOONGIGYVYNASTYCKPDPFIESHFLRVPVNDSPCEKILPWLDKSVDFIEKAKASNGC 239
DB 217 LDALGITALLINVSDDCPN-HFEGHYQYKCIPEVDNKHADISSWFMAIEYIDAVDKCRGR 275

QY 240 VLVHCLAGISRSATIAIAYIMKRMDSIDEAYRFVKEKRPITSPNPNLQGLDYEKKI 298
DB 276 VLVHCOAGISRSATICLAYIMWKRVLEBAFEFYKGRSRIISNFSFMQGLQFESGV 334

RESULT 10

US-09-702-705-805
Sequence 805, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 805
LENGTH: 394
TYPE: PRT
ORGANISM: Homo sapiens
US-09-702-705-805

Query Match 27.2%; Score 425; DB 4; Length 394;
Best Local Similarity 33.1%; Pred. No. 5.2e-38;
Matches 99; Conservative 62; Mismatches 116; Indels 22; Gaps 8;

QY 19 LESGTEKVLIDSRPVEYNTSHILEAININCSKLMKRLOQDVLITELI--QHSAXKH 76
DB 39 LPSSG-GKCLLDRCPLLAHSAGYILGSVNRCTIVRR-AKGSVSLQILPABEEVRAR 96
QY 77 VDIDCSQKVVVYDOSQDVASLSDCFLTVLGLKLEKSF--NSVHLIAGFAEFRCFPG 134
DB 97 LRSGLYAVIVYDERSPRASLRSDSTVSLVQALRRNARERTDILCKGGERFSSEYPE 156
QY 135 LCEGKSTL-----VPTCISOP-----CLPVA-----NIGPTRILPNLYGCGQDVLINKEL 179
DB 157 FCSKTKALAIIPPPVPASATPELDLDSCSGCTPLHDQEGVEILPFLYIGSAVHAARRDM 216
QY 180 IOONGIGYVYNASTYCKPDPFIESHFLRVPVNDSPCEKILPWLDKSVDFIEKAKASNGC 239
DB 217 LDALGITALLINVSDDCPN-HFEGHYQYKCIPEVDNKHADISSWFMAIEYIDAVDKCRGR 275
QY 240 VLVHCLAGISRSATIAIAYIMKRMDSIDEAYRFVKEKRPITSPNPNLQGLDYEKKI 298
DB 276 VLVHCOAGISRSATICLAYIMWKRVLEBAFEFYKGRSRIISNFSFMQGLQFESGV 334

RESULT 11

US-09-702-705-827
Sequence 827, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 827
LENGTH: 394
TYPE: PRN
ORGANISM: Homo sapiens
US-09-702-705-827

Query Match 27.2%; Score 425; DB 4; Length 394;
Best Local Similarity 33.4%; Pred. No. 5,2e-38;
Matches 100; Conservative 64; Mismatches 113; Indels 22; Gaps 9;

QY 19 LBSGTEKVLIDSRPEVNTSHILEAININCSKMKRRLOQDKVLTIELI--QHSAXKH 76
DB 39 LPSG-GKCLLDRCPLAHAGYILGSVNVRCNTIVRR-AGKSVLEQILPAEEVRRAR 96
QY 77 VVIDCSQKVVYVDOSSQDVASLSDCFLTVLGLKLEKSF--NSVHLAAGFAFSCFPG 134
DB 97 LBSGLYSAIVYDERPRAESLREDSTVSLVQALRNARERTDICKKGYSFSSSEYR 156
QY 135 LCEGKSTL-----VPTCISQ-----C-LPVANI-GPTRILPNULYGCQDVINKEL 179
DB 157 FCSKTKALAAIPPPVPASATEPLDLGSSCGTFLHDGCGFVELLPFLYLGSAVHAARRDM 216
QY 180 IQONGIGYVLNASYTCPKPDFPESHFLAVPNVDSFCEKILPWLDSVDFIEKAKSNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQKCIPEVDNKHADISSWFMEIETIDAVKOCGR 275
QY 240 VLVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPTISPNNFPLGQLLDYEKI 298
DB 276 VLVHCGAGISRSATICTAIYIMMKRVLEAEFEFVQKRSIISPNSFMGQLLQFSSQV 334

RESULT 12
US-09-736-457-805
Sequence 805, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darlick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 805
LENGTH: 394
TYPE: PRN
ORGANISM: Homo sapiens
US-09-736-457-805

Query Match 27.2%; Score 425; DB 4; Length 394;
Best Local Similarity 33.4%; Pred. No. 5,2e-38;
Matches 99; Conservative 62; Mismatches 116; Indels 22; Gaps 8;

QY 19 LBSGTEKVLIDSRPEVNTSHILEAININCSKMKRRLOQDKVLTIELI--QHSAXKH 76
DB 39 LPSG-GKCLLDRCPLAHAGYILGSVNVRCNTIVRR-AGKSVLEQILPAEEVRRAR 96

QY 77 VVIDCSQKVVYVDOSSQDVASLSDCFLTVLGLKLEKSF--NSVHLAAGFAFSCFPG 134
DB 97 LBSGLYSAIVYDERPRAESLREDSTVSLVQALRNARERTDICKKGYSFSSSEYR 156
QY 135 LCEGKSTL-----VPTCISQ-----C-LPVANI-GPTRILPNULYGCQDVINKEL 179
DB 157 FCSKTKALAAIPPPVPASATEPLDLGSSCGTFLHDGCGFVELLPFLYLGSAVHAARRDM 216
QY 180 IQONGIGYVLNASYTCPKPDFPESHFLAVPNVDSFCEKILPWLDSVDFIEKAKSNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQKCIPEVDNKHADISSWFMEIETIDAVKOCGR 275
QY 240 VLVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPTISPNNFPLGQLLDYEKI 298
DB 276 VLVHCGAGISRSATICTAIYIMMKRVLEAEFEFVQKRSIISPNSFMGQLLQFSSQV 334

RESULT 13
US-09-736-457-827
Sequence 827, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darlick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 827
LENGTH: 394
TYPE: PRN
ORGANISM: Homo sapiens
US-09-736-457-827

Query Match 27.2%; Score 425; DB 4; Length 394;
Best Local Similarity 33.4%; Pred. No. 5,2e-38;
Matches 100; Conservative 64; Mismatches 113; Indels 22; Gaps 9;

QY 19 LBSGTEKVLIDSRPEVNTSHILEAININCSKMKRRLOQDKVLTIELI--QHSAXKH 76
DB 39 LPSG-GKCLLDRCPLAHAGYILGSVNVRCNTIVRR-AGKSVLEQILPAEEVRRAR 96
QY 77 VVIDCSQKVVYVDOSSQDVASLSDCFLTVLGLKLEKSF--NSVHLAAGFAFSCFPG 134
DB 97 LBSGLYSAIVYDERPRAESLREDSTVSLVQALRNARERTDICKKGYSFSSSEYR 156
QY 135 LCEGKSTL-----VPTCISQ-----C-LPVANI-GPTRILPNULYGCQDVINKEL 179
DB 157 FCSKTKALAAIPPPVPASATEPLDLGSSCGTFLHDGCGFVELLPFLYLGSAVHAARRDM 216
QY 180 IQONGIGYVLNASYTCPKPDFPESHFLAVPNVDSFCEKILPWLDSVDFIEKAKSNGC 239
DB 217 LDALGITALLNVSSDCPN-HFEGHYQKCIPEVDNKHADISSWFMEIETIDAVKOCGR 275
QY 240 VLVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKEKPTISPNNFPLGQLLDYEKI 298
DB 276 VLVHCGAGISRSATICTAIYIMMKRVLEAEFEFVQKRSIISPNSFMGQLLQFSSQV 334

RESULT 14
US-09-614-124B-805
Sequence 805, Application US/09614124B

```

* Patent No. 6630574
* GENERAL INFORMATION:
* APPLICANT: Wang, Tonglong
* APPLICANT: Bangur, Chaitanya S.
* APPLICANT: Lodes, Michael A.
* APPLICANT: Fanger, Gary
* APPLICANT: Vedvick, Tom
* APPLICANT: Carter, Darrick
* APPLICANT: Retter, Marc
* APPLICANT: Mannion, Jane
* TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
* TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
* FILE REFERENCE: 210121.478C9
* CURRENT APPLICATION NUMBER: US/09/614,124B
* CURRENT FILING DATE: 2001-07-11
* NUMBER OF SEQ ID NOS: 1668
* SOFTWARE: FastSeq for Windows Version 3.0
* SEQ ID NO 805
* LENGTH: 394
* TYPE: PR1
* ORGANISM: Homo sapiens
* US-09-614-124B-805

```

Query Match	27.2%;	Score 425;	DB 4;	Length 394;
Best Local Similarity	33.1%;	Pred. No. 5.2e-38;		
Matches	99;	Conservative 62;	Mismatches 116;	Indels 22;
			Gaps	8

```

Qy      19 LESCTEKVLLIDSRPVEVNTYHTSHLEFALNINOSKLMERLOODVLLITELI--OHSKHK 76
Db      39 LPSC-GCCLLIDRRPLHSAAGITLHVSANRNTTVRR--AKGSVSEQLIPAEVEYRA 96

77 VDDICSQKVVVYVOQSSQDVASLSDCFLTVLLGKLEKSF--NSVHLLAGGFAEFSRCFPG 134
Qy      97 LRSGLSAVIVYDERSPRAELREDETVSLVQALRRRAERTDICLLKGGERPSSSEYPE 156
Db      135 LCBEKSTLV-----VPLCISQP-----CLFVA-----NIGPRLILPNYLGQGRDVLKEL 179
Qy      157 PCSKTKALAAIPPEVPSPASATEPDLDDSCSGTPLHDQSPVAILPELYLSAHTAARMD 216
Db      180 IQQNGIGVYNASTYCPKDPFIPESHFLRPVYNDSFCEKILFWLIDKSVDFEKKASNGC 239
Qy      217 LDALGTALANVSDCPN--HFEGHYQYKICIPAEVDHKADISSEMEFALEYIDAVKDRGR 275
Db      240 VLVNHCAGISASATIALYATKMDMSLDEAVRFEYERKPTSPFNFLIGOLLDYERKKI 298
Qy      276 VLVHCQGISASATICALYATKMKRYLBEAFPEFYQKRSITSPFMSMGLQLFESQV 334

```

RESULT 15
US-09-614-124B-827

```

; sequence 02. 6630574
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedrick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; OF TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 827
; LENGTH: 394
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-614-124B-827

```

Query Match	27.2%;	Score 425;	DB 4;	Length 394;
Best Local Similarity	33.4%;	Pred. No. 5.2e-38;		
Matches 100;	Conservative	64;	Mismatches 113;	Indels 22;
				Gaps 9;

[illegible]

Search completed: June 21, 2004, 13:25:44
Job time : 11.3061 secs

```

100; 27.2%; Score 425; DB 4; Length 394;
100; 33.4%; Pred. No.5.2e-38;
64; Mismatches 113; Indels 22; Gaps 9;

19 LBSGTEKVLILSRPVEYNTSHILEATINCSKLMKRLQGDVLTIELI--QHSATK 76
39 LPBG-GKCLILDRPFLASAGYILGSVVRNCTIVRR--AKGSVSLDELPAEEVPAR 96
77 VDDICSQKVVYDQSSQDVASLSSDFCLTVYLGLKEKS--NSVHLLAGFAPFEPKCPG 134
97 LRSGLSAVIVYDERSPRAESLRDSTVSLVVOALRRNARTDLCLEGGYERFSESYE 156
135 LCEGSKTL-----VPTCISQ-----C-LPVANI--GPTRLPLMYLGGCCRDYLINKEI 179
157 FCSKTVLAALPPVPVPSATEPLDLCSSCGTPLHDGSGPVETLLPPLYLGSAYHARRHM 216
180 IQONGYVYNASVYTPKPDFIPESHFLFVVPVNDSPCEKILPMLDSVDFLEKATASNGC 239
217 LDALGITALNVSSDCPN-HFEHGYQYKCIPEVDNKKADISWMFMAIYIDAIVXDRCGR 275
240 VLVHCLAGISRSATTAIVYIMKMDMSLDBAIVFVGEKAPTISPNNTFLGOLLDEYKKI 298
276 VLVHCGAGISRSATTCIAYIMKKKRYRLLEAEFEFVKORSIIISPNSFEGOLLQFESQV 334

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM proteoin - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 21:02:42 ; Search time 3544.67 Seconds
(Without alignment)

3692.750 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302

Perfect score: 1564

Sequence: 1 MAHEMGTGIVTERLVALL.....PNNFPLGOLDYKIKKNGT 302

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xjh
-Q/cgna2.1/USPTO.spool/US10029345/runcat.21062004.164146.29030/app.query.fasta.1.1294
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345 @CGN 1.1 6617 @runcat.21062004.164146.29030 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_ov.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_stb.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_mu.*
21: em_or.*
22: em_ov.*
23: em_ov.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_stb.*
28: em_un.*

29: em_vl.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pin.*
35: em_hcg_rtd.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hcgo_hum.*
40: em_hcgo_mus.*
41: em_hcgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1564	100.0	5450	6	AX482439 Sequence
2	1564	100.0	5450	6	AX482478 Sequence
3	1552	99.2	1998	6	AX260342 Sequence
4	1552	99.2	2071	6	AX921917 Sequence
5	1552	99.2	2118	6	AX099933 Sequence
6	1552	99.2	2732	6	AX180875 Sequence
7	1552	99.2	3059	6	AX278461 Sequence
8	1552	99.2	3104	6	AX405700 Sequence
9	1552	99.2	3284	6	BC042101 Homo sapi
10	1552	99.2	3496	6	AX441210 Sequence
11	1552	99.2	3521	9	AB052156 Homo sapi
12	1552	99.2	3544	6	AX260340 Sequence
13	1552	99.2	3566	9	AF506796 Homo sapi
14	1552	99.2	3766	6	AX374994 Sequence
15	1552	99.2	4790	6	BD171157 Novel gen
16	1552	99.2	4790	6	BD183422 Sequence
17	1552	99.2	4790	6	AB051467 Homo sapi
18	1545	98.8	2102	6	AX713989 Sequence
19	1545	98.8	2102	6	AX055973 Homo sapi
20	1544	98.7	5111	6	AX482372 Sequence
21	1517	97.0	4874	10	AB052157 Mus muscu
22	1517	97.0	4943	10	BC059232 Mus muscu
23	1517	97.0	4975	10	BC057321 Mus muscu
24	1514	96.8	4992	10	AF345951 Mus muscu
25	1504.5	96.2	2200	6	AX921915 Sequence
26	1492	95.4	2756	6	AX482444 Sequence
27	1375.5	87.9	4026	10	AF345953 Mus muscu
28	1312	83.9	2807	9	BC031643 Homo sapi
29	1246	79.7	749	6	AX868469 Sequence
30	1246	79.7	749	6	BD148531 Primer fo
31	1194.5	76.4	1935	9	AY038927 Homo sapi
32	1194.5	76.4	3332	6	AX441229 Sequence
33	1172.5	75.0	4827	10	AF345952 Mus muscu
34	1034	66.1	3861	10	AF345954 Mus muscu
35	1019.5	65.2	3231	5	BC063941 Dario rer
36	989	63.2	2264	9	BC045110 Homo sapi
37	988	63.2	2377	6	AR255944 Sequence
38	988	63.2	2377	9	HSU27193 Human prote
39	985	63.0	2453	6	A59888 Sequence 6
40	985	63.0	2453	10	WANTYPIGN
41	985	63.0	2600	10	BC052705 Mus muscu
42	943	60.3	2476	6	AX552119 Sequence
43	934.5	59.8	2415	6	A59887 Sequence 5
44	858.5	54.9	183797	9	AC022400 Homo sapi
45	858.5	54.9	198218	2	AC037447 Homo sapi

RESULT 1

ALIGNMENTS

AX482439
LOCUS AX482439 5450 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 108 from Patent WO02057460.
ACCESSION AX482439
VERSION AX482439.1 GI:22316984
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Todderud,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G.,
Suchard,S., Banas,D., Bassolino,D., Feder,J., Krystek,S.,
Mcatee,P., Mintier,G., Siemers,N., Jackson,D.G. and Ramathanan,C.
TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 108 25-JUL-2002;
SQUIBB BRISTOL MYERS CO (US)
FEATURES
Source
Location/Qualifiers
1..5450
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
538..2535
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD44451.1"
/db_xref="GI:22316985"
/translation="MAHEMIGTOIVTERLVALLSSTGTEKVLIDSRPFVEYNTSHILE
ATINCSKLMKRLQODKVLITELI,QHSKHKVDIDCSQKVVVDQSSQDVASISDPC
RTVLVLGKESFNSVHLIAGFAFSRCPGLCEKSTLTPCTISQPLPVANIQPT
RLPNVLGECQORDVANKELIQONGIGYVLAAYTCPKDPIDESHFLVFNDAFCEK
ILPWLDSVDFLEKASNGCVLHICLAGISRSATIAIAYIMKMDMSIDBAYFVE
KPTISPNPFLGOLLIDYEKLIKNOTGASGPKSLKILHLEKPNPVPAYSGCKSR
TLPSPCDSATSEAAQORPVHPASVSPVPSQBLSDSLVVALSGHLISADLSD
SNKRSRSLDKSYSSASMAASLHGSSSDALRYKBSITLIDGNKLCQFSPVGE
LSEQTPERSPDKEASITPKUQTAKPSDSQSKRLHSVTSSSGTAQSLSLPLRSGS
VEDNHTSFLPGLSTSOQLTKSNGLGIKMHSDILAPQTSPLPILTSWSYATSSHF
YSASAIYGGASAYSAYSCSLPTCGDQYVSVRRQKPEDRADSRSSWHEESPFEKQK
RRSCOMERGEISIMENRSREELGKVGSSPSGSMELIENS"

ORIGIN
Alignment Scores:
Pred. No.: 6,58e-151 Length: 5450
Score: 1564.00 Matches: 302
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x AX482439 (1-5450)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 538 ATGGCCCATGAGATGATGGAACCTCAATTCTTACTGAGAGGTTGGGCTCTGCTGAA 597
QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluIleAsnThrSer 40
DB 598 AGTGAACGGAAGGCTCTCTAATGATGACCGGCATTGTGGGAATCAATACATCC 657
QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleLeuMetCysValGluLeuGln 60
DB 658 CACATTTGGAGGCCATTAAATATCACTGCTCCCAAGCTTATGAGCGCAAGTTGCAACAG 717
QY 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
DB 718 GACAAAGTTAATTACAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 777
QY 81 CysSerGlnIleValIleValIleAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 778 TGCAAGTCAGAGGTTGATTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 837
QY 101 AspCysPheLeuThrValLeuLeuGlyIleLeuGluIleSerPheAsnSerValHisLeu 120

DB 838 GACTGTTTTCACCTGATCTTCTGGTAAACTGGAGAAAGACTTCACTGTTCACTGG 897
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGluIleSer 140
DB 898 CTTGCAAGTGGGTTTCTGCACTTCTCTGTTGTTCCCTGGCTCTGTGAAAGGAAATCC 957
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 958 ACTCTAGCTCCCTACCTCACTTCTCAGCTTGTCTTACTCTGTGCCAATTTGGGCAACC 1017
QY 161 ArgIleLeuProAsnLeuIleGlyValLeuAsnAlaSerThrCysProIleProAspPhe 180
DB 1018 GAAATCTTCCCAATCTTAACTTGTGGTCCAGCAGATCTCTCAACAGAGGCTGATA 1077
QY 181 GlnGlnAsnGlyIleGlyValLeuAsnAlaSerThrCysProIleProAspPhe 200
DB 1078 CAGCAGAAATGGAGTTGGTTATGCTTAAAGCCAGCTATCTGTCCAAAGCTTGACTTT 1137
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
DB 1138 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGTAAGACAGCTTTTGTGAGAAATTTTG 1197
QY 221 ProThrLeuAspIleSerValAspPheIleGluValAlaIleValIleValIleMet 240
DB 1198 CCGTGGTTGACAAATCATGATGATTCATTGAGAAAGCAAAAGCTTCAATGATGTGTT 1257
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
DB 1258 CTAGTGACATGTTTACCTGGAGATCTCCGCTCGGCACCATGCTTATGCTTACATCAG 1317
QY 261 LysArgMetAspMetSerLeuAspGluAlaIleThrPheValIleGluIleAspProThr 280
DB 1318 AAGAGATGACATGCTTTTAAATGAAAGCTTCAAGATTTGTGAAGAAAAAGACCTACT 1377
QY 281 IleSerProAsnPheAsnPheLeuGlyIleLeuLeuAspIleGluIleValIleAsn 300
DB 1378 ATATCTCCAAATCTTCAATTTTCTGGCCCAACTCTGACTATGAGAAAGATTAAGAC 1437
QY 301 GlnThr 302
DB 1438 CAGACT 1443
RESULT 2
AX482478
LOCUS AX482478 5450 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 147 from Patent WO02057460.
ACCESSION AX482478
VERSION AX482478.1 GI:22316999
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Todderud,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G.,
Suchard,S., Banas,D., Bassolino,D., Feder,J., Krystek,S.,
Mcatee,P., Mintier,G., Siemers,N., Jackson,D.G. and Ramathanan,C.
TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 147 25-JUL-2002;
SQUIBB BRISTOL MYERS CO (US)
FEATURES
Source
Location/Qualifiers
1..5450
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
538..2535
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD44453.1"
/db_xref="GI:22317000"
/translation="MAHEMIGTOIVTERLVALLSSTGTEKVLIDSRPFVEYNTSHILE

ATINCSKIMKRLQODKVLITELIOHSAKHVIDCSOKVWVVDOSODVASISDC
 FLTVLGLKLEKSPNSHYHLAAGFAERSRCPGLCEBKSTLPVPCISQPCPLANTCPT
 RILPNLYDQDRVNLKELIQONGIGVFLVASTTCRKPPIRPSHRLPVPNDSECEK
 ILPMIDKSVDFLEKAKASNGCVLHCLAGISRSATLAIYIMKRMMSIDEARFKE
 KRPITSPNPNFQOLIDYEKIKNOTGASPKSKLHLLEKNEBVSPLVSEGGQSE
 TPLSPCADSATSEAGORPVHPASVSPSVOPSLLEBSPLOALSGHLSADRLED
 LSKLRSPSLDIXSVSASMAARHSPSSERDLHYRPSRTLLDCTNKLCPSPVQE
 LSEQTPETSPDKREASIPKKLQOTRAPSDEOSRDLHVRTPSSSTARRSLSPLRSGS
 VEDNHTSPFLGISTSQOHLTKTSAGLGLKMSHDILAPOTSTPSTLSMTYFATESHS
 YSASAIYGSASAYSASCSQLPTCGDQVSVRRORRPSDRSRSRMSHESPEKQFK
 RRSQCMFBSIWSNRSREELKVGSGSFSGSMELIEVS"

ORIGIN

Alignment Scores:
 Pred. No.: 6,58e-151 Length: 5450
 Score: 1564.00 Matches: 302
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AX482478 (1-5450)

QY 1 MetAlaHigluMetilegIYThrgInileValThrgIuArgLeuValAlaLeuLeuGlu 20
 DB 538 ATGGCCCATGAGTGAATGGAACCTCAATTTGTTACTGAGAGTGGCTGCTGCTGGA 537
 QY 21 SerGIYThrgIuLyValLeuLeuileAapSerArgProheValGluIYraAnthrSer 40
 DB 598 AGTGAAGCGAAGAAAGTGGCTGTAATGATAGCCGGCATTTGTGGAATACAAATACATCC 657
 QY 41 HisIleleuGluAlaIleAenIleAenCySerIyLeuMetIyBarGLeuGInGIn 60
 DB 658 CACATTTTGAAGACCATTAATATCACTGCTCAAGCTTATGAAGCAAGGTGCAACAG 717
 QY 61 AspIyValLeuIleThrgIuLeuileGInHisSerAlaIyHisIyValAapIleAap 80
 DB 718 GACAAAGTGTAAATTACAGAGCTCATTCACATTCACGAAACATTAAGTTGACATTGAT 777
 QY 81 CysSerGInIyValValIYraSpGInSerSerGInAapValAlaSerLeuSerSer 100
 DB 778 TGCAGTCAGAGGTTGATGTTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 837
 QY 101 AspCyPheLeuThrValIleuLeuGlyIyLeuGluIySerPheAenSerValHisLeu 120
 DB 838 GACTGTTTCTCACTGACTTCTGGGTAACTGGAAGAGCTTCACTGTTCACTG 897
 QY 121 LeuAlaGlyIyPheAlaGluPheSerArgCyPheProGlyLeuCyGluGlyIySer 140
 DB 898 CTTCGAGGTGGTGTGCTGAGTCTCTCGTTGTTTCCCTGACCTCTGTGAAAGAAATCC 957
 QY 141 ThrLeuValProThrCySileSerGInProCySleuProValAlaAenIleGlyProThr 160
 DB 958 ACTCTAGTCCCTACCTGCTCACTTTCAGCTTGTGCTTACCTGTGCAACATTTGGGCCAAC 1017
 QY 161 ArgIleleuProAenLeuIYleuGlyCyGInArgAapValIleuAenIyGluLeuIle 180
 DB 1018 CGAATCTTCCCAATCTTATCTTGTGCTGCAAGAGATGCTCTCAACAAAGAGCTGATA 1077
 QY 181 GluGlnAenGlyIleGlyIYraValIleuAenAlaSerIYThrCySProLyPheProAapPhe 200
 DB 1078 CAGCAGATGGAGTGTATGTTAAATGCGAGCTTAACCTGTCCAAAGCCTGACTTT 1137
 QY 201 IleProGluSerHisPheLeuArgValProValAapAapSerPheCyGluIyIleLeu 220
 DB 1138 ATCCCGAGTCTATTTCTCGCGGTGCTGCTGTAATACAGCTTTTGTGAAAAATTTTG 1197
 QY 221 ProTTrpLeuAapIySerValAapPheIleGluIyValAlaIyAaSerAenGlyCyVal 240
 DB 1198 CCGGTGTGAGCAAAATCAGATGATTTCAATGAGAAACAAAGCCTCAATGATGTT 1257
 QY 241 LeuValHisCySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIYrIleMet 260

DB 1258 CTAGTCACTGTTTAGCTGGATCTCCGCTCCGCCACCATGCTATCCCTATCATG 1317
 QY 261 IyValGMeAspMeSerIleuSerGluAlaIYraGpPheValIyGluIyAaGpProThr 280
 DB 1318 AAGAGATGAGAAATGCTTTTAAATGAGAGCTTACAGATTTGTGAAGAAAGAACTTACT 1377
 QY 281 IleSerProAapPheAenPheLeuGlyIyGInIleuLeuAapIYrGluIySerIleIyAa 300
 DB 1378 ATATCTCAAACTTCAATTTTCTGGGCCAATCTGACTATGAGAAAGATTAAGAAC 1437
 QY 301 GInThr 302
 DB 1438 CAGACT 1443

RESULT 3
 AX260342
 LOCUS AX260342 1998 bp DNA linear PAT 26-OCT-2001
 DEFINITION Sequence 3 from Patent WO0173059.
 ACCESSION AX260342
 VERSION AX260342.1 GI:16509305
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 Meyers, R.A.
 38692 and 21117: dual specificity phosphatase molecules and uses
 therefor
 Patent: WO 0173059-A 3 04-OCT-2001;
 Millennium Pharmaceuticals, Inc. (US)
 JOURNAL
 FEATURES
 source Location/Qualifiers
 1..1998
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

Alignment Scores:
 Pred. No.: 3.19e-150 Length: 1998
 Score: 1582.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 1
 Query Match: 99.23% Indels: 0
 DB: 6 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AX260342 (1-1998)

QY 1 MetAlaHigluMetilegIYThrgInileValThrgIuArgLeuValAlaLeuLeuGlu 20
 DB 1 ATGGCCCATGAGTGAATGGAACCTCAATTTGTTACTGAGAGTGGCTGCTGCTGGA 60
 QY 21 SerGIYThrgIuLyValLeuLeuileAapSerArgProPheValGluIYraAnthrSer 40
 DB 61 AGTGAAGCGAAGAAAGTGGCTGTAATGATAGCCGGCATTTGTGGAATACAAATACATCC 120
 QY 41 HisIleleuGluAlaIleAenIleAenCySerIyLeuMetIyBarGLeuGInGIn 60
 DB 121 CACATTTTGAAGCCATTAATATCACTGCTCAACCTTAATGAAGCAAGGTGCAACAG 180
 QY 61 AspIyValLeuIleThrgIuLeuileGInHisSerAlaIyHisIyValAapIleAap 80
 DB 181 GACAAAGTGTAAATTACAGAGCTCATTCACAGCGAAACATTAAGTTGACATTGAT 240
 QY 81 CysSerGInIyValValIYraSpGInSerSerGInAapValAlaSerLeuSerSer 100
 DB 241 TGCAGTCAGAGGTTGATGTTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
 QY 101 AspCyPheLeuThrValIleuLeuGlyIyLeuGluIySerPheAenSerValHisLeu 120
 DB 301 GACTGTTTCTCACTGACTTCTGGGTAACTGGAAGAGCTTCACTGTTCACTG 360
 QY 121 LeuAlaGlyIyPheAlaGluPheSerArgCyPheProGlyLeuCyGluGlyIySer 140

Accession	Source	Length	Score	Identity	Positives	Negatives	Conserved	Mismatches	Indels	Gaps
Db										
361	CTTGAGAGTGGGTTTGGTCAAGTTCTCTGTTGTTTCCCTGGCCCTCTGGAAGGAATTC	420								
Qy	141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaIleAsnIleGlyProThr	160								
Db	421 ACTCTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACCTGTGGCAACATTGGCCCAAC	480								
Qy	161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnIleArgAspValLeuAsnIleGlyLeuIle	180								
Db	481 CGAATTCCTCCCAATCTTATCTTGGCTGCGCCAGCAAGATGCTCCCAAGAGAGCTGATG	540								
Qy	181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProIysProAspPhe	200								
Db	541 CAGCAGATGGGATTTGGTTATGTGTAAATGCCAGCAATCTGTCCAAAGCTGACTTT	600								
Qy	201 IleProIleSerHisPheLeuIleArgValProValAsnAspSerPheCysGlnIysIleLeu	220								
Db	601 ATCCCCAGATCTCAATTTCTGCGGTGCGCCGTGAATGACAGCTTTGTGTGAGAAATTTTG	660								
Qy	221 ProTyrPheAspIysSerValAspPheIleGlyIleValAlaValAsnIleArgValIleVal	240								
Db	661 CCGGTGTGGACAAATCAGTAGATTTCTTATGAAAGCAAAAGCTTCCAAATGATGTGT	720								
Qy	241 LeuValHisCysLeuAlaGlyIleSerArgSerValThrIleAlaIleAlaTyrIleMet	260								
Db	721 CTATGTGACATGTTTAACTGTGGAGTCTCCCGCTCCGCCACATCGCTATCCCTCAATCATG	780								
Qy	261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGlyIleValArgProThr	280								
Db	781 AAGAGATGGACATGCTTATAGTAGAAGCTTACAGATTTGTGAAAGAAAGAAAGACCTACT	840								
Qy	281 IleSerProAsnPheAsnPheIleGlyIleLeuLeuAspTyrGlyIleValIleValAsn	300								
Db	841 AATATCTCCAAATCTCAATTTCTGGGCGCAACTCTGACTATGAGAGAAAGATTAAAGAC	900								
Qy	301 GlnThr 302									
Db	901 CAGACT 906									
RESULT 4										
AX921917										
LOCUS	AX921917	2071 bp	DNA	linear	PAT 18-DEC-2003					
DEFINITION	Sequence 257 from Patent WO02068649.									
ACCESSION	AX921917									
VERSION	AX921917.1	GI:40215410								
KEYWORDS										
SOURCE	homo sapiens (human)									
ORGANISM	homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.									
JOURNAL	1									
FEATURES	Patent: WO 02068649-A 257 06-SEP-2002;									
source	Cutagen Corporation (US)									
	Location/Qualifiers									
	1..2071									
	/organism="Homo sapiens"									
	/mol_type="unassigned DNA"									
	/db_xref="taxon:9606"									
ORIGIN										
Alignment Scores:										
Pred. No.:	3,34e-150	Length:	2071							
Score:	1552.00	Matches:	300							
Percent Similarity:	99.67%	Conservative:	1							
Best Local Similarity:	99.34%	Mismatches:	0							
Query Match:	99.23%	Indels:	0							
DB:	6	Gaps:	0							

[illegible]

REFERENCE 1
AUTHORS Yue, H., Tang, Y. T., Bandman, O., Hillman, J. L., Baughn, M. R.,
Azizmai, Y. and Lu, D. A.
TITLE Protein phosphatase and kinase proteins
JOURNAL Patent: WO 012004-A 15 22-MAR-2001;
Incyte Genomics, Inc. (US)
FEATURES
source 1. 2118
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID NO: 1234795CB1"
ORIGIN
Alignment Scores:
Pred. No.: 3,44e-150 Length: 2118
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 6 Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x AX099933 (1-2118)
QY 1 MetAlAHieglumetiieglythrglnilevalThrgluarphleuValAlaleuLeuglu 20
DB 25 ATGGCCCATGAGATGATTGGAACTCAATGTTACTGAGGGTGGCTGCTGGAA 84
QY 21 SerGIYThrgluYsValleuLeuileaspSerArpProPhaValglurYraenThrSer 40
DB 85 AGTGAACGGAAGAAAGTGGCTGCTAATGATGACGGCCATTGGGATATCAATACATCC 144
QY 41 HieileugluAlaleanilleanleasnCyserLyseuMeuetyararagLeugln 60
DB 145 CACATTTTGAACCCATTATATCACTGCTCCAGCTTATGAAGCAAGTTGCAACAG 204
QY 61 AspLyseValleuileThrgluLeuileglnHieSerAlaleYsValleuLeu 80
DB 205 GACAAAGTGTATTTACAGAGCTCATTCACATTCACGCAACATAGGTTGACATTGAT 264
QY 81 CySerGlnlyeValValValValTyraapGlnSerSerGlnAspValAlaseLeuSerSer 100
DB 265 TGCAGTCAGAGGTTGATTGATTCAGATCAAGCTCCCAAGATGTTGCTCTCTTCA 324
QY 101 AspCyPhelLeuThrValleuLeuGlyLyseuGluYsSerPhaenSerValHleu 120
DB 325 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGCTTCAACTGTTCACTG 384
QY 121 LeuAlAGlylphelaAgluPheserArpCyPhaProglYeuCyseGluYlyser 140
DB 385 CTTCAGAGTGGTGTGCTGAGTCTCTCGTGTTCCTGCTGCTGAGGAAATCC 444
QY 141 ThleuValProThrCySileSerGlnProCyLeuProValAlaleanillegYProThr 160
DB 445 ACTGTCGCTCACTGCACTTCTGAGCTTGTCTACTGCTGCAACATGGGCCAAC 504
QY 161 ArgileuProAsnleuTyLeuGlyCyseGlnarGspValleuAnlyseGluLeu 180
DB 505 CGAATCTTCCCAATCTTATCTTGGCTGCCAGAGATGCTCTCAACAGAGAGCTGATG 564
QY 181 GluGlnAsnGlylleglyYrValleuAsnAlaserTyThrCySProlySProAspPhe 200
DB 565 CACAGAGATGGATGGTATGTTAATGACGAGCAATCCTGTCMAAGCCCTGATTT 624
QY 201 IleProGlnSerHiePhelAsnArgValProValAsnAspSerPheCyseGluYsileu 220
DB 625 ATCCCGAGCTCATTTCTCGCGGTGCTGATGATGACAGCTTTGTGGAATTTTG 684
QY 221 ProTrieuAspLyseSerValAspPheilegYlyeAlaleYsAlaseRanGlyCyVal 240
DB 685 CCGGTGTTGACAAATCAATGATTTTCATTTGAGAAAGCAAGCTCCCAATGATGTGTT 744
QY 241 LeuValHieCyseuAlaglylIseArGserAlathrlleAlaleYlyleuMet 260

DB 745 CTAGTCACTGTTTAGCTGGATCTCCGCTCGGCACCATGCTATCGCTCATCATG 804
QY 261 LyseArMetAspMeSerleuAspGluAlaTyraGPhaValleuGluYsArpProThr 280
DB 805 AAGAGATGATGATGCTTTCATGATGATGATGATGATGATGATGATGATGATGATG 864
QY 281 IleSerProAsnPhenAsnPhelLeuGlylileuLeuAspTyGluYsYsileYsAsn 300
DB 865 ATATCTCCAAACTTCAATTTCTGGCCCACTCTGACTATGAGAGAGATTAAAGAAC 924
QY 301 GlnThr 302
DB 925 CAGACT 930
RESULT 6
AX180875 2732 bp DNA linear PAT 06-AUG-2001
LOCUS
DEFINITION Sequence 2 from Patent WO0146394.
ACCESSION AX180875
VERSION AX180875.1 GI:15132703
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Plozman, G. D., Martinez, R., Whyte, D., Manning, G., Sudarshanam, S.,
Hill, R. J. and Flanagan, P.
TITLE Mammalian protein phosphatases
JOURNAL Patent: WO 0146394-A 2 28-JUN-2001,
Sugen, Inc. (US)
FEATURES
source 1. 2732
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.75e-150 Length: 2732
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 6 Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x AX180875 (1-2732)
QY 1 MetAlAHieglumetiieglythrglnilevalThrgluarphleuValAlaleuLeuglu 20
DB 538 ATGGCCCATGAGATGATTGGAACTCAATGTTACTGAGAGGTTGGCTGCTGGAA 597
QY 21 SerGIYThrgluYsValleuLeuileaspSerArpProPhaValglurYraenThrSer 40
DB 598 AGTGAACGGAAGAAAGTGGCTGCTAATGATGATGCGGCCATTGTGAAATACATACATCC 657
QY 41 HieileugluAlaleanilleanleasnCyserLyseuMeuetyararagLeugln 60
DB 658 CACATTTTGAAGCCATTATATCACTGCTCCAACTTATGAGAGGAGTTGCAACAG 717
QY 61 AspLyseValleuileThrgluLeuileglnHieSerAlaleYsValleuLeu 80
DB 718 GACAAAGTGTATTTACAGAGCTCACTGACATTCAGCAACATAGAGTTGACATTGAT 777
QY 81 CySerGlnlyeValValValValTyraapGlnSerSerGlnAspValAlaseLeuSerSer 100
DB 778 TGCAGTCAGAGGTTGATTGATTCAGATCAAGCTCCCAAGATGTTGCTCTCTTCA 837
QY 101 AspCyPhelLeuThrValleuLeuGlyLyseuGluYsSerPhaenSerValHleu 120
DB 838 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGCTTCAACTGTTCACTG 897

QY	121	LeuAlaGlyValPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer	140
Db	898	CTTGCAAGGTGGGTGGTGGAGTCTCTCGTTGTTTCCCTGGCTCTGTGAAGAAAAATCC	957
QY	141	ThrLeuValProThrCysAlleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	958	ACTCTAATCCCTTACCTGGACATTTTCAGCCCTGGTTAACCCTGGTCAACATTTGGGCCAACCC	1017
QY	161	ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuIle	180
Db	1018	CGAATCTTCCCAATCTTTATCTTGGCTGCCAGACAGATGTCTTCAACAGAGAGCTGATG	1077
QY	181	GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe	200
Db	1078	CACCAACAAATGGGATTTGGTATGTGTAATGCACAGCAATACCTGTCCAAACCTGACTTT	1137
QY	201	IleProGlnSerThrAspLeuArgValProValAsnAspSerPheCysGluLysIleLeu	220
Db	1138	ATCCCCGAGCTCATTTCTCGGTGTGCTCGTGAATGACACTTTTGTGAAGAAAATTTTG	1197
QY	221	ProTyrLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsnGlyCysVal	240
Db	1198	CCGTGTGTGACAAATCACTAATTTTCAATTTGAAGAAAGCAAGGCTCCAAATGATGTGTT	1257
QY	241	LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet	260
Db	1258	CTAAGTGCACCTGTTAGTGGGATGCCGCTCCGACACCAATCCGTATACGCTTACATATG	1317
QY	261	LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr	280
Db	1318	AMGAGAGATGGACATCTTTTAGATGAACTTACAGATTTGTGAAAGAAAGAAAAGCCTACT	1377
QY	281	IleSerProAsnPheAsnPheLeuGlyValLeuLeuAspTyrGlyLysIleLysAsn	300
Db	1378	ATATCTCAAACTTCAATTTTCTGGGCCAATCCTGGACATAGAGAAGATTTAAGAAC	1437
QY	301	GlnThr	302
Db	1438	CAGACT	1443

RESULT 7				
LOCUS	AX278461			
DEFINITION	AX278461	3059 bp	DNA	linear
ACCESSION	AX278461			PAT 02-NOV-2001
VERSION	AX278461.1			
KEYWORDS		GI:16605915		
SOURCE	.			
ORGANISM	Homo sapiens (human)			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
JOURNAL	Duecker,K.			
FEATURES	Identification of a dual specificity phosphatase: dupsp-10			
source	Patent: WO 0177340-A-1 18-OCT-2001; MERCK PATENT GmbH (DE) location/Qualifiers 1..3059			

```

:***** PRESENT GATHER (UBZ)
UBRS
source      Location/Qualifiers
            1. 3059
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"
CDS
            127_ 2124

```

```

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
127.
/feature="unassigned protein product"
/codon_start=1
/protein_id="CAD10545.1"
/db_xref="GI:16605916"
/translation="MAHEMTEGALIVTERIVALLSEGTGVKLLIDSRPREVNTSHILE
AIIINQSGELMKRFLQDDQVLTIGELIDSKAKHYDIDCGKVVYQSSDVAISLSDC
FLVATGLLEISFNSFVQVILGALGAEISRCFPLGCEKSTLVPTCTISQCLPVAIIGPT
RIILPMSVIVGORDGRIKELNQNGITGVYNAASNTCEKDPFIPESHFLVPPNDSFCR
LILPMLDSVDEIKAKNSCTVAILQGLAISGASATIALAIYIKRNDKSIDEAAYFVE
KRLPISINENFIDQLLDYERKINQGLAGSGPSKTKLHLKEPNEPVVAIVEGSGQFE

```

TRLEPQCADSTSEAGORPHNPASVSVSVOPSLLEDBEVLQALSGLHISADRLLED
 SNKLRSPSLDIKSVSNLWQASLHGFSSSEDLLEYKXKSTJLDGNKNCQFSPVQDE
 LSEDTPEPSPDEKASIPKSLQATRPBDSQEDRHLYRTKSTJLDSGAQSLSLPLHRSSG
 VEDNYHTSFLFGLSTSOOHLTKSAGLKLGMHNSIILADPSTPELSTSNWPATSSHP
 YSAIAIYGGASVASYVCSOLPTCGDGVYVRRRQKRSRDLSTFRHNEHSPEFKOPK
 RRSQMGFGEIMSENRNREBELKVGVSQSSFSQGMELTEVS"

Alignment Scores:	
pred. No.:	5,48e-150
Score:	1552.00
Percent Similarity:	99.67%
Best Local Similarity:	99.34%
Query Match:	99.23%
DB:	6
Gap:	0
Length:	305g
Matches:	300
Conservative:	1
Mismatches:	5
Indels:	0
Gap:	0

US-10-029-345A-109_COPY 1_302 (1-302) x AX278461 (1-3059)

QY	1	MeclahHsglmeCllleglYThgInlleValThrgluArgleValAlaleuLeuglu	20
Db	127	ATGGCCATGAGAGATTGGAAACCAATTGTTACTGAGAGGTGGCGCTGCTGGAA	186
QY	21	SerglYThrgluysValleuLeuIlleAspSerArpProhpeValgluyrAsnThSer	40
Db	187	AGTGAACGGAAAAAGTGGCTGCTAATTGATAGCCGCACTTGTGGAAATCAATACATCC	246
QY	41	HistleuGlualalleaenlleasnYsSerlySleuMetlysrArgleugIn	60
Db	247	CACATTTTGAAGCCATAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCCAACG	306
QY	61	AsplysValleuIleThrgluLeuIlleglnHisSerAlaHisAlaAspIlleap	80
Db	307	GACAAATGTTAATTACAGACTCATCCAGATTCCAGAAACATAGGTGCATTTGAT	366
QY	81	CysSerGlnlysValValValTyraSpGlnSerSerGlnAspValAlaSerleuSerAr	100
Db	367	TGCAGTCAGAGAGGTGATGATTACGATCAAAAGCTCCCAAGATGTTGCTCTCTTTTA	426
QY	101	AspCyPheleuthrValleuLeuglylylsleugluysSerPheasnSerValHisleu	120
Db	427	GACTGTTTTCTCACTGACTCTTGGGTAAACGTGAAGAGCTTCACTCTGTACCTG	486
QY	121	LeuAlaGlyGlyPheAlaGlnPheSerArgCySpPheProglyLeuCysegluGlyysSer	140
Db	487	CTTGACAGGGGGTGTGCTAGATTCCTGCTGCTTTTCCCTGGCGCTGTGAAGAAAAATCC	546
QY	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIlleglyProThr	160
Db	547	ACTCTAGTCCCTAAGCTGACATTTCTCAGCGTTGCTTACCTGTTGCCAAATTGGCCAAACC	606
QY	161	AsglIleleuProanLeuTyrlleuglyCyseglInArgaspValleuasnlysgluLeuIlle	180
Db	607	CGAATTCCTCCCAATCTTATCTTGGCGCCAGACAGATGTCCTCAACAAGAGCTGATG	666
QY	181	GlnGlnAsnGlylleglYTrValleuasnAlaSerTyrrhCysProlyProasphe	200
Db	667	CAGCAGAAATGGGATTGGTTATGTGTTAATCCAGCAATACCTGTCCAAAGCCGACATT	726
QY	201	IleProGlnSerHisPheLeuArgValProValAsnaspSerPheCysegluysIleleu	220
Db	727	ATCCCGGAGTCTCATTTCTGCGGTGGCTGTGAATGACGCTTTGTGAAGAAATTTTG	786
QY	221	ProThrleuAspIlySerValAspPheIllegluysAlaIysAlaSerAsnGlyCysVal	240
Db	787	CCGGGTGTGGCAAAATCAAGATTTCATAGAGAAACAAAGCCCTCAATGATGATGTT	846
QY	241	LeuValHisCysLeuAlaGlylIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet	260
Db	847	CTAGTGCACGTGTTAGCTGGGATTCCTCCGCTCCGCCACCATCGCTATCGCTCAATCATG	906
QY	261	LysArgMetAspMetSerleuAspGluAlaTyrrArgPheVallysgluysArgProThr	280
Db	907	AMGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAATACCTTACT	966

QY 281 115SerProAsnPhenAsnPhenLeuGlyGlnLeuLeuAspTyrGlnuyslyyileysAsn 300
 Db 967 ATATCTCCAACTTCAATTTCTGGGCGCAACTCTCGACTATGAGAAAGATTAAAGAAC 1026
 QY 301 GlnThr 302
 Db 1027 CAGACT 1032
 RESULT 8
 AX405700 3104 bp DNA linear PAT 14-JUN-2002
 LOCUS Sequence 115 from Patent WO0222660.
 DEFINITION AX405700
 ACCESSION AX405700.1 GI:21438839
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Tang, Y.-T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.-A., Ren, F.,
 Xue, A.-J., Yang, Y., Wehrman, T., and Dirmann, R. T.
 TITLE Novel nucleic acids and polypeptides
 JOURNAL Patent: WO 0222660-A 115 21-MAR-2002;
 HYSEQ, INC. (US)
 FEATURES
 source Location/Qualifiers
 1..3104
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 159..2156
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD34807.1"
 /db_xref="GI:21438840"
 /db_xref="REBASE:CAD34807"
 /translation="MAHMTGQIVTERVALLESGTEKYLIDSRPFVEVNTSHLE
 AINNSKLMKRLQDDKVLITTELIOHSAHKVDIDSCQVYVYDDSDVASLSDC
 PLVLGKLEKSFNSVHLAAGFAEFSRCFGLCEGKSTLPTICISQCLPVANIDPT
 RILPNDYLGQDVLNKEMLQONGIGVLAASNTCPKPDIPSHLRVAVNDSFCEK
 ILPMLDSVDFIEKAKASNGCVLHCLAGISRSATTAIATIMKRMMSIDEAYRFXE
 KRPTISPNFNFGLQLDYEKKIKNGTASGPKSKLLHLEKNEVPVAVSEGGQSE
 TPISPCDASATSEAGORPVHPASVSPVOPSLLEDSPLVQALSGILSADRED
 SNKLRKSFSLDIDKSVTSASMASLHGFSSSEDALEYRKSTLLDGTNKLQCPSPQIE
 LSEQTPETSPDKEEASIPKXLTQPARPSDSKRLHSVRSSTGTAQRLSLSPHRSISF
 VEDNHTSPFGLSTSQOHLTKSAGLTKGMSHDIILAPOTSTPLTSSWTFATSSHF
 YSASAIYGSASYSAYSCSQLPTCGDOQVYVRROKRPEDRADRRSMHEERSPREKQFK
 RRSQCMFEGSINSENRSRELKVKGSQSSFSGSMELIEVS"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.59e-150 Length: 3104
 Score: 1552.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 1
 Query Match: 99.23% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-345A-109_copy_1_302 (1-302) x AX405700 (1-3104)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuLeuGln 20
 Db 159 ATGGCCCATGATGATGATGAACTCAAAATGTTACTGAGAGTTGAGGCTCTGCTGGAA 218
 QY 21 SerGlyThrGlnuysValLeuLeuIleAspSerArgProPheValGlnuysThrSer 40
 Db 219 AGTGGAAAGGAAAGAGCTGCTGATGATGATGAGCCGCAATTTGTGCAATCAATACATCC 278
 QY 41 HisIleuGlnuysValIleAsnIleAsnCyseSerIleuysLeuysValArgArgLeuGln 60
 Db 279 CACATTTGGAGAGCATTAATATCAACTGCCCAAGCTTATGAAAGCAAGGTTGCAACAG 338

QY 61 AspLeuValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
 Db 339 GACAAAGTGTATATTCAGAGCTCATCCAGCAATTCAGCGAAATAGATTGACATTGAT 398
 QY 81 CyseSerGlnuysValValValIleAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 Db 399 TGCAGTCAGAGAGGTTGATGATTCAGATCAAGCTCCCAAGATGTTCCCTCTCTTCA 458
 QY 101 AspCysePheLeuThrValLeuLeuGlnuysIleuysLeuysSerPheLeuSerValIle 120
 Db 459 GACTGTTTTCTCACTGACTCTCGGGTAAACGGAAGAGCTTCAACTGTTCACCTG 518
 QY 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysePheProGlyLeuCyseGlnuysIle 140
 Db 519 CTTCGACAGTGGTTCCTGATGATTCCTGCTGTTCCCTGCTCTGTGAAGGAAAATCC 578
 QY 141 ThrLeuValProThrCyseIleSerGlnProCyseLeuProValAlaAsnIleGlyProThr 160
 Db 579 ACTCTAGTCCCTACCTGACATTCCTGACCTTGTCTGCTTACCTGCAACATTGGGCCAAC 638
 QY 161 ArgIleLeuProAsnLeuIleuysIleuysIleuysIleuysIleuysIleuysIle 180
 Db 639 CGAATCTTCCCAATCTTATATCTTGGCTCGCAGCGAGATGCTCTCAACAGAGCTGATG 698
 QY 181 GlnGlnAsnGlyIleGlyTyrValIleuysAlaSerTyrThrCyseProIleProAspPhe 200
 Db 699 CAGCAGAAATGGATGTTGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 758
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCyseGlnuysIleLeu 220
 Db 759 ATCCCGAGATCATTTCTCGCGTGTGCTGTAAGACAGCTTGTGTGAGAAAATTTTG 818
 QY 221 ProThrLeuAspIleSerValAspPheIleGlnuysAlaIleAsnAsnGlyCyseVal 240
 Db 819 CCCTGTTGTGACAAACAGTAGATTTTCATTGAGAAAGCAAAACCTCCAAATGATGTGT 878
 QY 241 LeuValHisCyseLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 Db 879 CTAGTCACTGTTTACCTGGATCTCCGCTCCGCCCAACATCGCTTACATCATG 938
 QY 261 LysArgMetAspMetSerLeuAspGlnuysIleuysIleuysIleuysIleuysIleuys 280
 Db 939 AAGAGATGAGACATGCTTAAATGAGAGCTTACAGATTGTGAAAAGAAAAGACTACT 998
 QY 281 115SerProAsnPhenAsnPhenLeuGlyGlnLeuLeuAspTyrGlnuyslyyileysAsn 300
 Db 999 ATATCTCCCAACTTCAATTTCTGGGCGCAACTCTCGACTATGAGAAAGATTAAAGAAC 1058
 QY 301 GlnThr 302
 Db 1059 CAGACT 1064
 RESULT 9
 BC042101 3284 bp mRNA linear PRI 07-OCT-2003
 LOCUS Homo sapiens dual specificity phosphatase 16, mRNA (cDNA clone
 DEFINITION MGC:50665 IMAGE:4400399), complete cds.
 ACCESSION BC042101 GI:27469788
 VERSION BC042101.1
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Strausberg, R. L., Feingold, B. A., Grouse, L. H., Derge, J. G.,
 Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D.,
 Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K.,
 Hopkins, R. F., Jordan, H., Moore, T., Ma, S. I., Wang, J., Hsieh, P.,
 Diatchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L.,
 Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L.,
 Scheetz, T. E., Brownstein, M. J., Umed, T. B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S. S., Loguettano, N. A., Peters, G. J.,

Abrahamson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hally, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 3284)
12477932
22388257

JOURNAL MEDLINE PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Strausberg, R.
Direct Submission
Submitted (23-DEC-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdexaxil@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 88 Row: a Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: Genomescan gene
prediction.

FEATURES
source
location/Qualifiers
1..3284
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:50665 IMAGE:4400399"
/tissue_type="Duodenal, adenocarcinoma"
/clone_id="NIH_MGC_88"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..3284
/gene="DUSP16"
/note="Synonyms: KIAA1700, MKP-7, MKP7"
/db_xref="LocusID:80824"
/db_xref="MIM:607175"
332..3230
/codon_start=1
/product="DUSP16 protein"
/protein_id="AAH42101.1"
/db_xref="GI:27469789"
/translat="LocusID:80824"
/translation="MAHEMIGTQIVTERLVALLESGETKYLIDSRPFVYNTSHILE
AININGSKMRRLQODKVLITELIOHSAKXKVIDDSOKVYVDSOAVLASNPC
FLTVLIGKTESFNSVHLIAGFAFSPRCVCEGSTIIVPTCTISQPCPVANIGPT
RIIPNVLGCRDVIANKELMQNGITGVLANSTCPDPDPSPSHFLRVVNSFCCK
ILPWLDSVDFIEKASNGCVLHICLAGISRSATIAIAYIMKMDSLDEARFVE
KLPTISPNFNLGLDYEEKIKXQGTASGPKSLKLHLEKPEVPVAVSEGOQSE
TPLSPCADSATSEAGQRPVHPASVPSVPLEDDPILVQALSGHLISADRLDSNK
LKRFSFLDIKSVSASMAASLHGFSSSEDALEYKSTLDGNTKLCOPSPOLSE
OTPEPTSPDKERASTPKLQTAIRSDSGSKRLHVSRTSSGTAORSILSPHRSQVSD
NYHRSPLFGIISTSQOHLTKSAGGLKXMHSDILAPQTSPLTSLSMTFATSSHFPISA
SAIYGSASTAISCSQPLTCGDVIVSVRRRQKPSRADSRRSWHSESPPEKQPKRS

CDS
gene

misc_feature
380..730
/note="RHOD: Rhodanese Homology Domain"
/db_xref="CCD:smart00450"
misc_feature
803..1222
/note="USPC: Region: Dual specificity phosphatase,
catalytic domain. Ser/Thr and Tyr protein phosphatases.
The enzyme's tertiary fold is highly similar to that of
tyrosine-specific phosphatases, except for a 'recognition'
region"
/db_xref="CCD:pfam00782"

ORIGIN
Alignment Scores:
Pred. No.: 6e-150 Length: 3284
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
Gaps: 0
DB:
US-10-029-345a-109_copy_1_302 (1-302) x BC042101 (1-3284)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
Db 332 ATGGCCCATGAGATGATGGAGTCAATGTTACTGAGAGGTTGGTGGCTGTGCGAA 391
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 392 AGTGAAGCGAAAGAGCGTGAATGTAAGCGCGCATTTGTGAATTAACAATCAATCC 451
QY 41 HisIleLeuGlnIleAlaIleAsnIleAsnCyseSerLeuMetLysArgArgLeuGln 60
Db 452 CACATTTTGAACCATTTATATCACTGCTCAACCTTATGAAGGAGTTGCAACAG 511
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 512 GACAAAGTTTAAATTACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATGAT 571
QY 81 CysSerGlnLysValValValValYraAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 572 TGCAGTCAGAGGTGTGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 631
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db 632 GACGTGTTTCTCACTGATCTTGGGTTAACTGGAAGAAGCTTCAACTGTTCACCTG 691
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCyseGluGlyLysSer 140
Db 692 CTTCAGAGTGGGTTCTGCTGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 751
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 752 ACTCTGATCCCTACCTGACATTTCTGACCTTGTCTTCTGCTGCTGCTGCTGCTGCTGCT 811
QY 161 ArgIleLeuProAsnLeuThrLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
Db 812 CGAATCTTCCCATCTTATCTTGGCTGCGAGGAGATGCTCTCAACAAGAGCTGAG 871
QY 181 GlnGlnAsnGlyIleGlyLysValLeuAsnAlaSerThrCysProLysProAspPhe 200
Db 872 CAGCAAAAGTGAATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTAT 931
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
Db 932 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991
QY 221 ProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAsnGlyCysVal 240
Db 992 CCGTGTGTGAACAATCAGATTTTCATTTGAGAAAGCAAAAGCTTCAATGAGATGTGTT 1051
QY 241 LeuValHisCysLeuAlaGlyLysSerArgSerAlaThrIleAlaIleAlaIleAlaIleAla 260

Db 1052 CTAGTGCACGTGTAGCTGGAGATCTCCCGCTCCGCCACCATGCTATCGCTACATCATG 1111

Qy 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluLysArgProThr 280

Db 1112 AAGAGATGAGACATGCTTTAGATGAGAGCTTACAGATTGTGTGMAAAGAAAAAGACCTACT 1171

Qy 281 ILeSerProAspPheAspPheLeuGlyIleLeuLeuAspTyrGluLysIleLysAsn 300

Db 1172 ATATCTCCAAATCTCAATTTCTGGGCCAACCTCTGACTATGAGAAAGATTAAAGAAC 1231

Qy 301 GlnThr 302

Db 1232 CAGACT 1237

RESULT 10

AX441210 3496 bp DNA linear PAT 28-JUN-2002

LOCUS Sequence 1 from Patent WO226997.

DEFINITION AX441210

ACCESSION AX441210

VERSION AX441210.1 GI:21665766

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Luche, R.M. and Wei, B.

AUTHORS Dep-16 dual-specificity phosphatase

TITLE Patent: WO 0226997-A 1 04-APR-2002;

JOURNAL Cepcyr, Inc. (US)

FEATURES

source 1..3496

location/Qualifiers

1..3496 /organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 6.5e-150 Length: 3496

Score: 1552.00 Matches: 300

Percent Similarity: 99.67% Conservative: 1

Best Local Similarity: 99.34% Mismatches: 1

Query Match: 99.23% Indels: 0

DB: 6 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AX441210 (1-3496)

Qy 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20

Db 562 ATGGCCCATGAGATGATGAGACTCANAATTGTACTGAGAGGTGGCTGCTGCTGGA 621

Qy 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40

Db 622 AGTGGAAAGCGAAAGAGCGCTGCTAATTGATGAGCGGCACTTGTGGAATACATACATCC 681

Qy 41 HisIleLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 60

Db 682 CACATTTTGGAGCCATTATATATCACTGCTCCACGCTTATGAGCGAAGGTTGCAACAG 741

Qy 61 AspLeuValLeuLeuIleThrGluLeuIleGlnHisSerAlaLysHisIleValAspIleAsp 80

Db 742 GACAAAGTGTATATACAGAGCTACCTCCAGCATTCAGCGAAACATNAGGTGACATTGAT 801

Qy 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100

Db 802 TGCAGTCAGAGAGTTGATGATTGATCGATCAAGCTCCCAAGATTTGCTCTCTTCA 861

Qy 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120

Db 862 GACTGTTTCTCACTGACTCTGGGTTAACTGGAAGAGCTTCAACTCTGTTCACTCG 921

Qy 121 LeuValGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140

Db 922 CTTCGAGGAGGGTTTGCTGATGTTCTCTGTTGTTTCCCTGCGCTGTGAGAGAAAAATCC 981

Qy 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160

Db 982 ACTCTAGTCCCTTACCTGATCTTCTCAAGCTTCTTACCTGTTCCAAACCTTGGCCAAAC 1041

Qy 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGluLeu 180

Db 1042 CCAATTTCTCCAAATCTTATCTTGGCTGCCAGCGAGATGCTCTCAACAGAGCTGATG 1101

Qy 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200

Db 1102 CAGCAGAAATGGATGGTATGTTAATGACAGAAATACCTGCTCAAAAGCTGACTTT 1161

Qy 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220

Db 1162 ATCCCGGATCTCATTTCTGCTGGTGGCTCTGAAATGACAGCTTTGTGAGAAATTTTG 1221

Qy 221 ProTProLysAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240

Db 1222 CCGTGGTGGACAAATCACTAATTCATTGAGAAAGCAAAAGCTTCCATAGATGTGTT 1281

Qy 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260

Db 1282 CTAGTGCACGTGTAGCTGGATCTCCCGCTCCGCCACCATGCTATGCTACATCATG 1341

Qy 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluLysArgProThr 280

Db 1342 AAGAGATGAGACATGCTTTAGATGAGAGCTTACAGATTGTGTGMAAAGAAAAAGACCTACT 1401

Qy 281 ILeSerProAspPheAspPheLeuGlyIleLeuLeuAspTyrGluLysIleLysAsn 300

Db 1402 ATATCTCCAAATCTCAATTTCTGGGCCAACCTCTGACTATGAGAAAGATTAAAGAAC 1461

Qy 301 GlnThr 302

Db 1462 CAGACT 1467

RESULT 11

AB052156 3521 bp mRNA linear PRI 18-OCT-2001

LOCUS Homo sapiens MKP-7 mRNA for MAPK phosphatase-7, complete cds.

DEFINITION AB052156

ACCESSION AB052156

VERSION AB052156.1 GI:13548676

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Masuda, K., Shima, H., Watanabe, M. and Kikuchi, K.

AUTHORS MKP-7, a novel mitogen-activated protein kinase phosphatase, functions as a shuttle protein

TITLE J. Biol. Chem. 276 (42), 39002-39011 (2001)

JOURNAL MEDLINE 21486429

PUBMED 11489891

REFERENCE 2 (bases 1 to 3521)

AUTHORS Masuda, K., Shima, H. and Kikuchi, K.

TITLE Direct Substitution

JOURNAL Submitted (05-DEC-2000) Kouhei Masuda, Institute for Genetic Medicine, Division of Biochemical Oncology and Immunology, Kita-ku, Kita 5-Jou nishi 1-7-Cyume, Sapporo, Hokkaido 060-0815, Japan (E-mail: kou@imm.hokudai.ac.jp, Tel: 81-11-706-5536, Fax: 81-11-707-6839)

FEATURES

source 1..3521

location/Qualifiers

1..3521 /organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="12"

/cell_line="Jurkat"

1..3521 /gene="MKP-7"

CDS

564..2561
/gene="MKP-7"
/codon_start=1
/product="MAPK phosphatase-7"
/protein_id="BAB40814.1"
/db_xref="GI:13548677"
/translation="MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILE
AININGSKLMKRLQODKVLITELIGHSAHKVLDIOSQVYVDDSSODVASLSDC
FLTVLLGKLEKSFNSVHLIAGFAEPRCPGLCEGKSTIVPTCISQPCLVANIGPT
RIPLNLYGCGRDVNLKELMOQNGITGVLANSTCEPDPFIPESHFLRVYVDSFCEK
ILPWLDSVDFIEKAKASNGCVLHCLAGISRSATLAIAYIMKMDSLDEAYRVE
KRPTISPNFNLGQLDYERKIKNQTASGPKSLKLHLHEKPEPVPAYSEGGQSE
TPLEPCADSATSEAGQRPVHPASVSPVPSLDESPVLQALSGHLISADRLD
SNKIKSFSLDIKSVASASMAASLHGFSSSEALHEYKSTPLDGNKLCOFSPVE
LSBOETPSDPKEEASIPKKLQTPARSDSQSKRLHSVRSSTSGTARSLSLPHRSGS
VEDNYHTSFLPGLSTSOQLTKSAGLKGKMSDILAPQSTPSLTSSWYFATESSHF
YSASAIYGGASYSAYSCSQLPTCGDVYSVRRRQKPSDRADSRRSWHESPEKQPK
RRSCMEFGESIMSENRSRRELKGVGSQSSFSGSMELIYS"

ORIGIN

Alignment Scores:
Pred. No.: 6,55e-150 Length: 3521
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AB052156 (1-3521)

OY 1 MetAlahisglumellleglythrglnilevalthrgluargleuvalaleuleu 20
DB 564 ATGGCCCATGAGATGATTGGAACCAATTGTTACTGAGAGTTGGTGGCTCTGGAA 623
OY 21 SerGlythrgluysvalleuleuileaspseratgprophevalglutyrasnthser 40
DB 624 AGTGAACGGAAAAAGTCCTGCTAATTGATAGCCGGCATTTGGGAATACAAATACATCC 683
OY 41 HisileuenglualialeasnilleasnCyserislyseumellysarxargleuengl 60
DB 684 CACATTTGGAAAGCCATTAAATCACTGCTCCAGCTTATGAAAGCGAAGTTGCCACAG 743
OY 61 AsplysValleuilethrgluuileuileglnhisseralalyshislyvalasp 80
DB 744 GACAAAGTTAATTAACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 803
OY 81 CysSerGlnlysvaivalvaltyrAspGlnSerSerGlnaapvalalaserleuSer 100
DB 804 TGCAGTCAGAAAGTTGTGATTTCATCAAAAGCTCCAAAGATGTTGCTCTCTCTTCA 863
OY 101 AspCysPheleuThrValleuleuengllylsleuengllysserPheAsnSerValhisleu 120
DB 864 GACTGTTTTCACCTGCTCTCTGGGTAACTCGAAGAAAGACTTCACTCTGTTCACTG 923
OY 121 LeuAlaglygylphealaglypheserargCysPheProgllyleuCyegllylysser 140
DB 924 CTTCGAGGTGGGTTGCTGAGATTCTCTGTTGTTTCCCTGGGCTCTGGAAGGAAATCC 983
OY 141 ThrleuValProthCysIleSerGlnProCysleuProvalAlaAsnIleGlyProthr 160
DB 984 ACTCTAGTCCTACCTGCAATTCCTGAGCTTGCTTACCTGTTGCCAATGGGCGAAACC 1043
OY 161 ArgIleleuProAsnleuTyrlleuGllyCysGlnaargapvalleuAsnlysglu 180
DB 1044 CGAATTCCTCCAACTTATCTTCTGCTGCCAGCAAGATGTCCTCAACAAGAGCTGATG 1103
OY 181 GlnGlnaenglylleglytyrValleuAsnAlaserlyrThrCysProlyspProaspPhe 200
DB 1104 CAGCAGATGGGATGTTAGTGTAAATGCAACATACCTGTCCAAGCTGACCTTT 1163
OY 201 IleProGlnSerHisPheleuargValaProvalaAsnAspSerPheCysGlnlyslleu 220
DB 1164 ATCCCGAGCTCATTTCTCTGGTGGCTGTGAATGACAGCTTTGTGAGAAATTTTG 1223

OY 221 ProTripleuAsplysSerValaaspheileglulysalalyalaseranglyCysVal 240
DB 1224 CCGTGTGTGACAAATCAGTAAATTTCTTGAAGAAAGCAAAAGCTCCATGATGATGTT 1283
OY 241 LeuValHisCysleuAlaglyllyeserargseralathrillealallealalyrilemet 260
DB 1284 CTAGTGACACTGTTTACCTGGATCTCCCGCTCCGCCACCACTTCCTATCGCTTACATCATG 1343
OY 261 LyArgMetAspMetSerleuAspGluAlaTyArgPheVallysgluysargProthr 280
DB 1344 AAAGAGATGACATGCTTTTGAATGAAGCTTAACAGATTGTGAAAAGAAAAAGACTTACT 1403
OY 281 IleSerProAsnPheAsnPheleuengllyglnleuAspTyrgluyslyllyeAsn 300
DB 1404 ATATCTCAAACTTCATTTTCTGGCGCAACTCCTGACATATAGAAAGAAATTAAGAAC 1463
OY 301 GlnThr 302
DB 1464 CAGACT 1469

RESULT 12
AX260340
LOCUS AX260340 3544 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 1 from Patent WO0173059.
ACCESSION AX260340
VERSION AX260340.1 GI:16509303
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Meyers, R.A.
TITLE 38692 and 21117: dual specificity phosphatase molecules and uses
JOURNAL therefor
PUBLISHED Patent: WO 0173059-A 1 04-OCT-2001.
Milestone Pharmaceuticals, Inc. (US)
FEATURES
source
1.3544
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
589..2586
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10218.1"
/db_xref="GI:16509304"
/translation="MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHILE
AININGSKLMKRLQODKVLITELIGHSAHKVLDIOSQVYVDDSSODVASLSDC
FLTVLLGKLEKSFNSVHLIAGFAEPRCPGLCEGKSTIVPTCISQPCLVANIGPT
RIPLNLYGCGRDVNLKELMOQNGITGVLANSTCEPDPFIPESHFLRVYVDSFCEK
ILPWLDSVDFIEKAKASNGCVLHCLAGISRSATLAIAYIMKMDSLDEAYRVE
KRPTISPNFNLGQLDYERKIKNQTASGPKSLKLHLHEKPEPVPAYSEGGQSE
TPLEPCADSATSEAGQRPVHPASVSPVPSLDESPVLQALSGHLISADRLD
SNKIKSFSLDIKSVASASMAASLHGFSSSEALHEYKSTPLDGNKLCOFSPVE
LSBOETPSDPKEEASIPKKLQTPARSDSQSKRLHSVRSSTSGTARSLSLPHRSGS
VEDNYHTSFLPGLSTSOQLTKSAGLKGKMSDILAPQSTPSLTSSWYFATESSHF
YSASAIYGGASYSAYSCSQLPTCGDVYSVRRRQKPSDRADSRRSWHESPEKQPK
RRSCMEFGESIMSENRSRRELKGVGSQSSFSGSMELIYS"

CDS
ORIGIN
Alignment Scores:
Pred. No.: 6,61e-150 Length: 3544
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AX260340 (1-3544)


```
QY      1 MetAlaHtGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db      589 ATGGCCCATGAGTGAATGGAACTCAAAATTGTTACTGAGAGTTGGTGGCTCTGCTGGAA 648
QY      21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluIlyrAanThrSer 40
Db      649 AGTGAACGGAAGAAAGTCTGCTAATGATGACCGGCACTTTGTGGAAATCAATATCACTCC 708
QY      41 HttIleLeuGluAlaIleAenIleAenCysSerIleLeuMetIleValArgLeuGln 60
Db      709 CACATTTTGAAGCCATTAAATATCACTGCTCAAGCTTATGAGAGCGAAGTTGCAACAG 768
QY      61 AspLeuValLeuIleThrGluLeuIleGlnHisSerAlaValHisValAspIleAsp 80
Db      769 GACAAAGTGTAAATTACAGAGCTCATCAGCATTCACGCAAACTAAGGTTGACATTGAT 828
QY      81 CysSerGlnValValValIlyrAspGlnSerSerGlnAspAlaIleSerLeuSerSer 100
Db      829 TGCAGTCAGAAAGGTGTAGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 888
QY      101 AspCysPheLeuThrValIleLeuLeuGlyValLeuGluIlyrSerPheAenSerValHisLeu 120
Db      889 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGCTTCACTCTGTTCACCTG 948
QY      121 LeuAlaGlyValPheAlaGluPheSerArgCysPheProGlyLeuGlyGluIlyrSer 140
Db      949 CTTCAGAGTGGGTTTGTGAGTCTCTGCTGTTGTTCCCTGGCTGTGGAAGAAATCC 1008
QY      141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAenIleGlyProThr 160
Db      1009 ACTCTAGTCCCTACCTGCACTTCTGACGCTTGTGCTTACTGTTGCAACATTGGGCCAAC 1068
QY      161 ArgIleLeuProAenLeuIlyrLeuGlyCysGlnArgAspValLeuAenIlyrGluLeu 180
Db      1069 CGAATCTTCTCCATCTTATCTTGGCTGCGCAGGAGATGCTCTCAACAAGAGCTGATG 1128
QY      181 GlnGlnAenGlyIleGlyIlyrValIleAenAlaSerIlyrThrCysProIlyrAspPhe 200
Db      1129 CAGCAGATGGAGTGGTATGTTGTTAAAGCCAGCAATCTGTTCCAAAGCCCTGACTTT 1188
QY      201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIlyrIle 220
Db      1189 ATCCCCAGCTCATTTCTCGCGTGGCTGCTGATATACGCTTTGTGGAATAATTTTG 1248
QY      221 ProTyrLeuAspIlyrSerValAspPheIleGluValAlaValAsnArgIlyrCysVal 240
Db      1249 CCGTGTGGTGAACAATACATGATTTTCATTGAGAAACAAAGCCTCCATGATGCTT 1308
QY      241 LeuValHisCysIleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIlyrIleMet 260
Db      1309 CTAGTGCACTGTTTACCTGGGATCTCCGCTCCGCCCAATCGCTATCGCTTACATCATG 1368
QY      261 LysArgMetAspMetSerLeuAspGluAlaIlyrArgPheValIlyrGluIlyrAspProThr 280
Db      1369 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAAACCTACT 1428
QY      281 IleSerProAenPheAenPheLeuGlyGlnLeuLeuAspTyrGluIlyrIleValAsn 300
Db      1429 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGAGATATGAGAAAGATTAGAAC 1488
QY      301 GlnThr 302
Db      1489 CAGACT 1494
```

```
RESULT 13
AF506796      3566 bp      mRNA      linear      PRI 30-OCT-2003
LOCUS      Homo sapiens dual specificity phosphatase 16 (DUSP16) mRNA,
DEFINITION      complete cdb.
ACCESSION      AF506796
VERSION      AF506796.1 GI:25573087
KEYWORDS      Homo sapiens (human)
```

```
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 3566)
AUTHORS      Hoornaert, I., Marynen, P., Gorle, J., Sciote, R. and Baens, M.
TITLE      MAPK phosphatase DUSP16/MKP-7, a candidate tumor suppressor for
chromosome region 12p12-13, reduces BCR-ABL-induced transformation
JOURNAL      Oncogene 22 (49), 7728-7736 (2003)
PUBMED      14586399
REFERENCE      2 (bases 1 to 3566)
AUTHORS      Hoornaert, I., Marynen, P. and Baens, M.
TITLE      Direct Submision
JOURNAL      Submitted (26-APR-2002) Department for Human Genetics-Flanders
Interuniversity Institute for Biotechnology (VIB), Katholieke
Universiteit Leuven, Herestraat 49, Leuven B-3000, Belgium
FEATURES
source
1..3566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p13-p12"
/issue_type="Done marrow"
1..3566
/gene="DUSP16"
/note="synonym: MKP7"
633..2630
/gene="DUSP16"
/EC_number="3.1.3.48"
/EC_number="3.1.3.16"
/codon_start=1
/product="dual specificity phosphatase 16"
/protein_id="PANT5120.1"
/db_xref="GI:25573087"
/translation="MAHEWIGTQIVTERLVALLSESTEKVLIDRSPFVEYNTSHILE
ALINCSKMKRRLQODKVLITELIOHSAKHVDIDSCQKVVYDQSDVASLSDC
FLTVLLGKLEKSPNSVHLIAGGPAERSRCPGLCEGKTLVPTCISQPLVANIQT
RLRNLVLCQGDVLYNKELMOONGCYVYNAANTCKPRPIESHRLRPVYDSFEK
ILPWLDSVDFLEKASNGCYLVHGLAETISATTAIVIMKRDMSIDAEVAYKE
KRPITSPNFFLGOLDYERKLIKNQDASGPSKLLHLERNEVPAVSEGGQKSE
TPLSPCADASAEAGCPRVHPASVPSVQPSLIEDSPVQALSGHLSLDRIED
SNLKRSPFLDIKSVYSASMAASLHGFSSSDALERYKPTLIDGTNLCQSPVOB
LSBQNPETSPOKEBASIPKLTOTAPRSDOSRSHLVRSSSGTAQORSLPLHREGS
VEDNHTSPLRELSTSQOHLTSAAGLGKMSDILAPOTSTPSTLSSYPRARESSH
YASATYGGASAYSQQLPTCGQVTSVRKQDPSDRAUSRRKRWHEBSPPEKQPK
RSCQMEFEGESIMSENRREBELKXGVSQSFSGSMELIEVS"
```

ORIGIN

```
Alignment Scores:
Pred. No.:      6,66e-150      Length:      3566
Score:      1552.00      Matches:      300
Percent Similarity:      99.67%      Conservative:      1
Best Local Similarity:      99.34%      Mismatches:      0
Query Match:      99.23%      Indels:      0
DB:      9      Gaps:      0
```

US-10-029-345A-109_COPY_1_302 (1-302) x AF506796 (1-3566)

```
QY      1 MetAlaHtGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db      633 ATGGCCCATGAGTGAATGGAACTCAAAATTGTTACTGAGAGTTGGTGGCTCTGCTGGAA 692
QY      21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluIlyrAanThrSer 40
Db      693 AGTGAACGGAAGAAAGTCTGCTAATGATGACCGGCACTTTGTGGAATCAATATCACTCC 752
QY      41 HttIleLeuGluAlaIleAenIleAenCysSerIleLeuMetIleValArgLeuGln 60
Db      753 CACATTTTGAAGCCATTAAATATCACTGCTCAAGCTTATGAGAGGTTGCAACAG 812
QY      61 AspLeuValLeuIleThrGluLeuIleGlnHisSerAlaValHisValAspIleAsp 80
Db      813 GACAAAGTGTAAATTACAGAGCTCATCAGCATTCACGCAAACTAAGGTTGACATTGAT 872
```


QY	81	CysSerGlnIuYsValValValIlyrAapGlnSerSerGlnAspValAlaSerLeuSerSer	100
Db	873	TGAGAGTCAGAAAGGTGTAGTTTACGATCAAAAGCTCCAAATGTTGCTCTCTCTTCA	932
QY	101	AspCysPheLeuThrValLeuLeuGlyLYrLeuGlnIuYsSerPheAsnSerValIhIsleu	120
Db	933	GACTGTTTTCTCACTGACTTCTGGGTAACCTGGAGAAAGGCTTCAACTCTGTTCACTTG	992
QY	121	LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGlnGlyLysSer	140
Db	993	CTTGCAAGTGGGTTTGGCTGAGTTCTCTCGTTGTTTCCCTGGCCTCTGTGGAAGAAATTC	1052
QY	141	ThrIleuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	1053	ACTCTAGTCCCTACCTCACTCATTTCTCAAGCTTGTCTTACTTGTCACATCTGGCCCAAC	1112
QY	161	ArgIleLeuProAsnLeuIlyrLeuGlyCysGlnArgAspValLeuAsnIlyrGluLeuIle	180
Db	1113	CGAATTTCTTCCCATCTTTATCTTGGCTGGCCAGAGATGTCCTCAACAAGAGAGCTGATG	1172
QY	181	GlnGlnAsnGlyIleGlyIlyrValLeuAsnAlaAseryThrCysProLysProAspPhe	200
Db	1173	CAGCAGATGGAGTGGTTGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTT	1232
QY	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIuYsIleu	220
Db	1233	ATCCCGAGTCTCAATTTCTTCGCTGCTGCTGTGATGACAGCTTTGTGGAAAAATTTTG	1292
QY	221	ProThrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal	240
Db	1293	CCGGTGTGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCAATGATGTGCTT	1352
QY	241	LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIlyrIleMet	260
Db	1353	CTAGTGCACTGTTTAGCTGGAGATCTCCGCTCCGCAACATCGCTATCGCTCAATCATG	1412
QY	261	LysArgMetAspMetSerLeuAspGlnAlaIlyrArgPheValLysGlnLysArgProThr	280
Db	1413	AAAGAGATGACACACTCTTTAGATGACAGCTTACAGATTGTGAAAAAGAAAAAGACTACT	1472
QY	281	IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIlyrGlnLysLysIleLysAsn	300
Db	1473	ATATCTCAAACTTCATTTTCTGGGCCAACTCTCGACTATGAGAAAGATTAAGAAC	1532
QY	301	GlnThr 302	
Db	1533	CAGACT 1538	
RESULT 14			
LOCUS	AX374994	3766 bp	DNA
DEFINITION	Sequence 17 from Patent WO0210363.		
ACCESSION	AX374994		
VERSION	AX374994.1	GI:19169826	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Wang, Y.-T., Elliott, V.-S., Ramkumar, J., Yao, M.-G., Burford, N.,		
	Tang, Y.-E., Stewart, E.-A., Gandhi, A.-R., Patterson, C., Lee, B.-A.,		
	Hafidli, A.-J., Lu, D.-A., Tribouley, C.-M., Griffin, J.-A., Baughn, M.-R.,		
	Yue, H., Warren, B.-A., Nguyen, D.-B. and Wallia, N.-K.		
TITLE	Protein phosphatases		
JOURNAL	Patent: WO 0210363-A 17 07-FEB-2002;		
	Incyte Genomics, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..3766		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		

ORIGIN	/note="Incyle ID No: 7480570CB1"
Alignment Scores:	7.14e-150
Pred. No.:	1552.00
Score:	99.67%
Percent Similarity:	99.34%
Best Local Similarity:	99.23%
Query Match:	6
DB:	Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x AK374994 (1-3766)	
QY	1 MetAlaHisGluMetIleGlyThrGlnIleValIleThrGlnArgLeuValAlaLeuGln 20
DB	538 ATGGCCCATGATGATGATTGGAATCTCAATATTGTTACTGAGAGGTTGGTGCTGCTGGAA 597
QY	21 SerGlyThrGlnValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
DB	598 AGTGGAAAGGAAAAAGTGTCTAATTGATAGCCGGCATTTGTGAAATACATATCATCTC 657
QY	41 HisIleLeuGlnAlaIleAsnIleAsnCysSerIleLeuMetIleValArgLeuGlnGln 60
DB	658 CACATTTTGGAAAGCCATTAATATCACTGCTCCAGCTTATGAAAGGAAAGCTTGCACAG 717
QY	61 AspIleValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
DB	718 GACAAAGTGTATTAATACAGAGCTCATTCACATTCACGGAACATAGAGTTGACATTGAT 777
QY	81 CysSerGlnValValIleValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB	778 TGCAGTCGAAAGGTTGTAGTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 837
QY	101 AspCysPheLeuThrValLeuLeuGlyIleLeuGlnIleCysSerPheAsnSerValHisLeu 120
DB	838 GACGTTTTTCCACTGTACTTCTGGGGTAACTGGAGAAAGACTTCAACTCTGTTCACCTG 897
QY	121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuGlyIleValSer 140
DB	898 CTTGCAAGTGGGTTTGGCTGAGTTCTCGTTGTTTCCCTGGCCCTGTGAAAGGAAATCC 957
QY	141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB	958 ACTCAGACCTCCACTGCTGATTTCCACAGCTTGCTTACCTGTGTGCAACATTGGGCCAAC 1017
QY	161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGluLeuIle 180
DB	1018 CGAATCTTCCCAACTTTATCTTGGCTGCCAGAGGATGCTCTCAACAAGAGACTGATG 1077
QY	181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProIleProAspPhe 200
DB	1078 CAGCAGATGGGATGGTATGTGTGTTAATCCAGCAATACCTGTCCAAAGCCCTGACCTT 1137
QY	201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleLeu 220
DB	1138 ATCCCGAGTCTCATTTCTCTCGCGGTGCTGTGAATGACAGCTTTGTGAGAAAATTTTG 1197
QY	221 ProThrLeuAspIleSerValAspPheIleGluIleValAlaValAlaSerAsnGlyCysVal 240
DB	1198 CCGTGTTGGACAAATCTGATGATTTCTTGTAGAAAGCAAAAGCTCCAAATGATGTGCT 1257
QY	241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
DB	1258 CTAAGGCACTGTTTGTGCTGGGATCTCCGCTCCGCAACATTCGCTATTCGCTCAATCATG 1317
QY	261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValIleGluIleValArgProThr 280
DB	1318 AAGAGAGATGAGACATCTCTTGTAGATGAACCTTACAGATTGTGAAAGAAAAGACCTTACT 1377
QY	281 IleSerProAsnPheAsnPheLeuGlyIleLeuLeuAspTyrGluIleValIleValAsn 300
DB	1378 ATATCTCCAAATCTTAATTTCTGGGCCAATCTCTGAGCTATGAGAAAGATTAAGAAC 1437

Qy 301 GlnThr 302
Db 1438 CAGACT 1443

RESULT 15
BD171157

LOCUS BD171157 4790 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel gene and protein encoded thereby.
ACCESSION BD171157
VERSION BD171157.1 GI:27876969
KEYWORDS WO 02052005-A/13.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 4790)
Ohtsuka, O., Nagase, T. and Nakajima, D.
TITLE Novel gene and protein encoded thereby
JOURNAL Patent: WO 02052005-A 13 04-JUL-2002;
KAZUSA DNA RESEARCH INSTITUTE, OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE
NAKAJIMA
COMMENT OS Homo sapiens (human)
PN WO 02052005-A/13
PD 04-JUL-2002
PF 20-DEC-2001 WO 2001JP011217
PR 22-DEC-2000 JP 00P 389742
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/12, C07K14/47
CC Novel gene and protein encoded thereby
FC Key location/Qualifiers
FT CDS location/Qualifiers (184).. (2178).

FEATURES
Source 1..4790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 9.69e-150 Length: 4790
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: Gaps: 6

US-10-029-345A-109_COPY_1_302 (1-302) x BD171157 (1-4790)

Qy 1 MetAlaHisGluMetCileGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 184 ATGGCCCATGAGATGATGGAACCAATTGTTACTGAGAGTTGGTGGCTCTGGCGAA 243

Qy 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 244 AGTGGAAAGGAAAGAGCTGCTAATGATGATGATGATGATGATGATGATGATGATGAT 303

Qy 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIleLeuLeuLeuLeuLeuLeuLeuLeu 60
Db 304 CACATTTTGGAGGACATTAATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 363

Qy 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleValHisIleValLeuLeu 80
Db 364 GACAAAGTTTAATTAACAGAGCTCATCAGCATTCAGCAACATTAAGTTGACATTGAT 423

Qy 81 CysSerGlnIleValIleValIleValIleValIleValIleValIleValIleValIleVal 100
Db 424 TGCAGTCAGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483

Qy 101 AspCysPheLeuThrValLeuLeuGlyIleLeuGluIleValSerPheAsnSerValHisLeu 120
Db 484 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAGAAAGAGCTTCACTGTTCACTTG 543

Qy 121 LeuAlaGlyIleGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyIleValSer 140

Db 544 CTTCGAGGTGGGTTCCTAGTTCCTCGTTGTTTCCCTGGCCTCTGTGAAGGAAATCC 603

Qy 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160

Db 604 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGCTTACCTGCTGCAACATTTGGCCAAACC 663

Qy 161 ArgIleLeuProAsnLeuIleGlyCysGlnArgAspValLeuAsnIleGlyLeuLeu 180

Db 664 CGAATTCCTCCCAATCTTATCTTGGCTGCGAGAGATGCTCTCAACAAAGAGCTGATG 723

Qy 181 GlnGlnAsnGlyIleGlyThrValLeuAsnAlaSerThrCysPheProIleProAspPhe 200

Db 724 CAGCAAAAGGAGATGCTTATGCTTAAATGCCAGCAAAATCCTGCAAAAGCTGACTTT 783

Qy 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIleVal 220

Db 784 ATCCCGAGTCTCATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 843

Qy 221 ProThrLeuAspIleSerValAspPheIleGlyValAlaIleValAsnGlyCysVal 240

Db 844 CCGTGGTGGACAAATCAGTATGATTCATTGAGAAAGCAAAAGCTTCAATGATGATGTT 903

Qy 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleValMet 260

Db 904 CTAGTCACTGTTTACCTGATCTCCGCTCCGCAACATGCTATCCCTTACATCATG 963

Qy 261 LysArgMetAspMetSerLeuAspGluAlaIleArgPheValIleGluIleValArgProThr 280

Db 964 AAGAGGATGACATGCTTATGATGAGCTTACAGATTGTGAAAGAAAAAGACTTACT 1023

Qy 281 IleSerProAsnPheAsnPheLeuGlyGlnIleLeuAspThrGluIleValIleValAsn 300

Db 1024 ATATCTCCAAACTTCAATTTCTGGGCAACTCTGACTATGAGAAAGATTAAGAAC 1083

Qy 301 GlnThr 302

Db 1084 CAGACT 1089

Search completed: June 22, 2004, 02:43:22
Job time : 3555.67 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 20:53:18 ; Search time 2662.97 Seconds
(without alignments)
3985.196 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMGTQVTERVALLE.....PNFPGQLDYEKKIKNGT 302

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xjh
-Q=/cgr2_1/USPTO.spool/US10029345/runat_21062004_164146_29039/app_query.fasta_1.1294
-DB=EST -QPM=fastap -SUPPIX=ret -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPM=prc -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US10029345_QCGN_1_1_3931_@runat_21062004_164146_29039 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST :
1: em_estda :
2: em_esthum :
3: em_estlin :
4: em_estmu :
5: em_estov :
6: em_estpl :
7: em_estro :
8: em_hic :
9: gb_est1 :
10: gb_est2 :
11: gb_hic :
12: gb_est3 :
13: gb_est4 :
14: gb_est5 :
15: em_estfun :
16: em_estom :
17: em_gss_hum :
18: em_gss_hiv :
19: em_gss_pln :
20: em_gss_vit :
21: em_gss_fun :
22: em_gss_mam :
23: em_gss_mus :
24: em_gss_pro :
25: em_gss_rtd :
26: em_gss_phg :
27: em_gss_vil :
28: gb_gss1 :

29: gb_gss2 :
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1369.5	87.6	898	13	BQ721265
2	1145.5	73.2	836	13	B1821804
3	1099	70.3	715	14	CF727177
4	1086.5	69.5	682	12	B1816954
5	1081	69.1	656	14	CF532917
6	1076	68.8	639	13	CF742387
7	1069	68.4	769	13	BU704078
8	1006	64.3	650	10	BF472046
9	989	63.2	2207	11	BC038231
10	978.5	62.6	924	14	CA474739
11	959	61.3	862	13	BQ770036
12	936	59.8	602	10	AM847425
13	931	59.5	595	10	AM847425
14	891	57.0	3325	11	AK035652
15	872	55.8	792	13	BU946569
16	853	54.5	946	12	B1857528
17	823.5	52.7	516	29	AY413655
18	804.5	51.4	516	29	AY413657
19	772.5	49.4	683	12	BM942528
20	770	49.2	619	29	CG583869
21	769	49.2	579	10	BF582366
22	762	48.7	512	10	BF815601
23	725	46.4	1060	12	BM546940
24	703	44.9	421	10	BE946884
25	697.5	44.6	1180	12	BM908762
26	695	44.4	404	10	BB816652
27	693	44.3	463	10	CG058779
28	675.5	43.2	412	9	AA879894
29	675	43.2	903	12	BG482429
30	665	42.5	442	14	CB606149
31	659.5	42.2	881	14	CF266082
32	645	41.2	402	29	CG499793
33	637	40.7	463	10	BF809325
34	635	40.6	667	14	CA394418
35	630.5	40.3	801	14	CK130134
36	615.5	39.4	685	14	CB165612
37	608	38.9	459	10	BF566061
38	597.5	38.2	403	10	BE004625
39	596.5	38.1	1148	12	BM806608
40	586	37.5	494	13	BQ318152
41	586	37.5	494	13	BQ318155
42	582	37.2	861	14	CA325142
43	577	36.9	635	10	BE377720
44	571	36.5	1043	12	BM563030
45	569	36.4	337	14	CB606296

ALIGNMENTS

RESULT 1
BQ721265
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

BQ721265 898 bp mRNA linear EST 16-JUL-2002
AGENCOURT_8294508 lupsk1_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6194455 5', mRNA sequence.
BQ721265
BQ721265.1 GI:21860162
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 898)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strusberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: Agencourt Bioscience Corporation
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LLM1359 row: 1 column: 08
 High quality sequence stop: 669.

FEATURES
 source
 1. .898
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6194455"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult" 16 yr"
 /lab_host="DH10B"
 /clone_lib="lupski_sympathetic_trunk"
 /note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
 NotI; Site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCAGCGCTCCG-3' and
 5'-GACTAGTTCAGTACGAGCGGCCGCCCT(15)-3'. Site selected >
 1 kb for average insert length 1.9 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine); available through Life
 Technologies."

ORIGIN

Alignment Scores:
 Pred. No.: 1,786-152 Length: 898
 Score: 1369.50 Matches: 279
 Percent Similarity: 95.27% Conservative: 3
 Best Local Similarity: 94.26% Mismatches: 12
 Query Match: 87.56% Indels: 4
 DB: 13 Gaps: 1

US-10-029-345a-109_COPY_1_302 (1-302) x BQ721265 (1-898)

QY 4 GtUmeTlIeGlyThrGlnIleValThrgIuTrgLeuValAlaLeuLeuGluSerGlyThr 23
 DB 14 GAGATGATGGAACTCAATTTCTTACTGAGAGTTGGTGGCTCTGCTGGAAAGTGGACG 73
 QY 24 GtUlvSVallLeuLeuIleAspSerTrgProPheValGluTyraThrSerHisIleLeu 43
 DB 74 GAAAGAGTGTCTAATGATAGCCGGCCATTGTGGAATACATATACATCCACATTTTG 133
 QY 44 GtUaIaIleAsnIleAsnCySerIleYsLeuMeTlYsAargArgLeuGlnGlnAspLysVal 63
 DB 134 GAAAGCATTAATATCACTCTCCAGAGTTTGAAGCCAAAGGTTGCAACGAGAAAGTG 193
 QY 64 LeuIleThrGluLeuIleGlnHisSerAlaIleYsHisIySVaIleAspIleAspCySerGln 83
 DB 194 TTAATTTACAGAGCTCATCCAGCATTCAGCAACATAGATGATGATGATGAGTTCAG 253
 QY 84 LysValIleValIleYsAspGlnSerSerGlnAspValAlaSerLeuSerSerAspCyPhe 103
 DB 254 AAGGTGTAGTTTACATCAAGCTCCCAAGATTTTCTCTCTCTTCCAGACTGTTT 313
 QY 104 LeuThrValLeuLeuGlyIleYsLeuGluYsSerPheAsnSerValHisLeuLeuAlaGly 123
 DB 314 CTCACGTGACTTCTGGGTAACGTGAGAAAGCTTCAACTCTGTTCACTGCTTGGAGGT 373
 QY 124 GlyPheAlaGluPheSerTrgCyPheProGlyLeuCyGluGlyIleYsSerThrLeuVal 143

DB 374 GGGTTTGCTGAGTCTCTCGTGTGTTCCCTGGCCCTGTGAGAGAAATCCACTCTAGTC 433
 QY 144 ProThrCySileSerGlnProCyLeuProValAlaAsnIleGlyProThrArgIleLeu 163
 DB 434 CCTACTGTCATTTCTGACCTCTTCTTACTCTGTTCGCAACATGGGCAACCCGAATCTT 493
 QY 164 ProAsnLeuTyrlLeuGlyCySglNArgAspValIleuAsnYsGluLeuIleGlnGlnAsn 183
 DB 494 CCCATCTTTTTCCTGGCTGCACGAGATGTCTCAACAAGAGGCTGATGACGAGCAT 553
 QY 184 GlyIleGlyTyraValIleAsnAlaSerTyrlThrCySprolyProAspPheIleProGlu 203
 DB 554 GGGATGGTTTATGTCTTAATGCCAGCAATACCTGTCCAAAGCCTGACTTATCCCGAG 613
 QY 204 SerHisPheLeuArgValProValAsnAspSerPheCySgluYsIleLeuProTyrPleu 223
 DB 614 TCTCATTTCTCGCGTGTGCGCTGTGAATGACAGCTTTGTAGAAATTTTTCGCGTGTG 672
 QY 224 AspIysSerValAspPheIleGlyValAlaYsAlaSerAsnGlyCySVaIleuValHis 243
 DB 673 GACAAATCAGTACATTTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTCTTAGTGAC 732
 QY 244 CySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlIleMetIysArgMet 263
 DB 733 TGTTTACTGGATCTCCCGCTCGC-ACCATGCTATGCTTCACTACATGAGAAAGATG 791
 QY 264 AspMetSer-LeuAspGluAlaTyrlArgPheValYsGluYsArgProThrIleSerPr 283
 DB 792 GACATGCTTTTGAAGAACTTACAGATTTTGGAAAGAAAGCAAGCACTACTATATCTCC 851
 QY 283 cAsnPheAsn--PheLeuGlyGlnLeuAspTyrlGluYsYs 297
 DB 852 AAAACTTCAATTTTCTTGCGCACTCTCGCATTTATGAAAAA 897

RESULT 2
 BI821804
 LOCUS 836 bp mRNA linear EST 04-OCT-2001
 DEFINITION 603035883F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176724 5',
 mRNA sequence.
 ACCESSION BI821804
 VERSION BI821804.1 GI:15933354
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strusberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LLM11440 row: f column: 21
 High quality sequence stop: 805.

FEATURES
 source

1. .836
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5176724"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 115"
 /note="Organ: pooled brain, lung, testis; Vector:
 PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is

oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	9.27e-126	Length:	836
Score:	1145.50	Matches:	244
Percent Similarity:	96.08%	Conservative:	1
Best Local Similarity:	95.69%	Mismatches:	8
Query Match:	73.24%	Indels:	7
DB:	12	Gaps:	0

US-10-029-345A-109_COPY_1_302 (1-302) x B1821804 (1-836)

```

QY 50 CysSerLysLeuMetLysArgArgLeuGlnGlnAspValLeuLeuThrGluLeuLeu
DB 2 TGTCTCCAGGCTTATGAAAGCGAGGTTGCAACAGACAGATGTTAATTACAGAGCTCATC 61
QY 70 GlnHisSerLysLysHisLysValAspLeaSpCysSerGlnLysValValValTyrAsp 89
DB 62 CAGCATTCAGCGAAGAACTAAGTTGACATTTGA-TGCAGTCAGAAAGGTTGAGTTAGAT 120
QY 90 GlnSerSerGlnAspValAlaSerLysSerSerAspCysPheLeuThrValLeuLeuGly 109
DB 121 CAAGAGCTCCCAAGATGT-GCCTCTCTCTCTTCAAGAGCTTTTCTCACTGACTTCCGGGT 179
QY 110 LysLeuGlnLysSerPheAsnSerValHisLeuLeuAlaGlyLysPheAlaGluPheSer 129
DB 180 AAATCTGAGAAAGAGCTTCAACTGTCACCTCTGACAGTGGTGGTGGTGGTCTCT 239
QY 130 ArgCysPheProGlyLeuCysGlyGlyLysSerThrLeuValProThrCysLiesGln 149
DB 240 CGTAGTTTCCCTGGCTCTGTGAAGAAATCCACTTACTCTCCCTACCTGATTTCTCAG 239
QY 150 ProCysLeuProValAlaAsnLleGlyProThrArgLleLeuProAsnLeuTyrLeuGly 169
DB 300 CTTTGGCTTACTTGGCCAACTTGGGCGCAACCGAATCTTCCCAATCTTATCTTGGC 359
QY 170 CysGlnArgAspValLeuAsnLysGluLeuLleGlnGlnAsnGlyLleGlyTyrValLeu 189
DB 360 TGCAGAGAGATGTCTTCAACAGAGAGCTGATGAGAGATGGATGGATGATGATG 419
QY 190 AsnAlaSerTyrThrCysProLysProAspPheLleProGluSerHisPheLeuArgVal 209
DB 420 AATGCCAGCAATACCTCTCCAAAGCCGACTTATCCCGAGTCTCATTTCCGCGGTG 479
QY 210 ProValAsnAspSerPheCysGlyLysLleLeuProTyrLeuAspLysSerValAspPhe 229
DB 480 CCGTGTGATGACAGCTTCTGTGAGAAATTTTCCCGTGGTGGACAAATCAGTAGATTTTC 539
QY 230 LLeGlnLysValAlaLysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyLieser 249
DB 540 ATTGAGAAACCAAAAGGCTCCAAATGAGATGTGT-CTAGTGCAGCTTTAGCTGGAGTCTCC 588
QY 250 ArgSerLysThrLleAlaLleAlaTyrLleMetLysArgMetAspMetSerLeuAspGlu 269
DB 599 CGCTCCGC-ACCATCGCTATCGC-TACATCATGAAAGAGATGAGACATGCTTTTAGTGA 656
QY 270 AlaTyrArgPheValLysGlu-LysArgProThrLieserProAsnPheAsnPheLeu-G 289
DB 657 GCTTACAGATTTGTGACAGAAAGAGACCTACTATATCTCCAAACTTCAATTTTCTGGG 716
QY 289 LysGlnLeuAspTyrGlnLysLysLysLysLysLysLysLysLysLysLysLysLys 302
DB 717 GCCAGCTCTGAGCTATGAGAAAGATTAAGAAACGAGACT 757

```

RESULT 3
CF727177 715 bp mRNA linear EST 09-OCT-2003

DEFINITION UI-M-HB0-ckj-3-09-0-UI.r1 NIH BMAP_HB0 Mus musculus cDNA clone
IMAGE:30548096 5', mRNA sequence.
ACCESSION CF727177
VERSION CF727177.1 GI:37601345
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 715)
AUTHORS NIH-MGC <http://mgi.mgi.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contract: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

seq primer: pyx-5.
Location/Qualifiers
1..715
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30548096"
/issue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1b="NIH BMAP_HB0"
/note="Organ: Eye; Vector: pyx-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTAATGAGT. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

FEATURES

source

1..715
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30548096"
/issue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1b="NIH BMAP_HB0"
/note="Organ: Eye; Vector: pyx-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTAATGAGT. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Alignment Scores:

Pred. No.:	2.58e-120	Length:	715
Score:	1099.00	Matches:	206
Percent Similarity:	98.15%	Conservative:	6
Best Local Similarity:	95.37%	Mismatches:	4
Query Match:	70.27%	Indels:	0
DB:	14	Gaps:	0

US-10-029-345A-109_COPY_1_302 (1-302) x CF727177 (1-715)

```

QY 87 ValTyrAspGlnSerSerGlnAspValAlaSerLysSerSerAspCysPheLeuThrVal 106
DB 2 GTTATGATCAAGAGTTCCAAAGATGTGTTCTGTGTGTCAGACAGCTTTTCTCATGTGA 61
QY 107 LeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeuLeuAlaGlyLysPheAla 126
DB 62 CTTCTGGTAACTGAGAGAGAGACTTCAACTGTCTCAGCTGTTTCAGGTGGCTTGGCT 121
QY 127 GluPheSerArgCysPheProGlyLeuCysGlyGlyLysSerThrLeuValProThrCys 146
DB 122 GAGTTCTCTGCTGTTGTTCCCTGGCTGTGTAAGAAATGCCACTGACTGCTTACTCTGC 181

```

```
QY 147 IISerGlnProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnIleu 166
Db 182 AATCTCAGCCTTCTGCTTACCTGTGCAACATGGGCCAAGCTCCAAATCTCCCAATCTC 241
QY 167 TTTLeuGlyCysGlnArgAspValLeuAsnIleGlyLeuIleGlnIleAsnGlyTleGly 166
Db 242 TATCTGTGGCTGCGACGAGATGCTCTCAACAAGACCTGACGACCAAGATGGATTGGC 301
QY 187 TTTValIleAsnAlaSerTyrThrCysProLysProAspPheIleProGlnSerHisPhe 206
Db 302 TATGCTTTAAATGCGACGAATACCTGTCCAAAGCCTGATCTTCAATCCGAAATCTCACTTC 361
QY 207 LeuArgValProValAsnAspSerPheCysGlnIleuIleLeuProTyrLeuAspLysSer 226
Db 362 CTGCGAGTGGCTGTGAATGACAGCTTTTGTGAGAAATCTTACCATGTTGGACAGTCT 421
QY 227 ValAspPheIleGlnIleValAlaValAlaSerAsnGlyCysValLeuValHisCysLeuAla 246
Db 422 GTGGATTTCAATGGAAAGCAAAAGCCTCCAAATGGCTGTGTGCTTATCCACTGCTTAGCT 481
QY 247 G1YIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLysArgMetAspMetSer 266
Db 482 GGGATCTCTGCTCGCGCAATATGCTATTGCTTACATCAATGAGAGAGATGACATGTCT 541
QY 267 LeuAspGlnAlaTyrArgPheValLysGlnLysArgProThrIleSerProAsnPheAsn 286
Db 542 CTGATATAGGCTTACACATTTGTGAAGAAAGAAAGACCTTCTATCTCGAATTTTAT 601
QY 287 PheLeuGlyGlnLeuLeuAspTyrGlyLysLysIleLeuAsnGlnThr 302
Db 602 TTTATGGCCACATCATGACATATGAGAGACATTAATTAACCACT 649
```

RESULT 4

```
BI816954/c 682 bp mRNA linear EST 10-OCT-2001
LOCUS imagec.10.2000/s1z410bdf41.x1 Soares_NPBMC Homo sapiens cDNA
DEFINITION clone IMAGE:4140798 3', mRNA sequence.
ACCESSION BI816954
VERSION BI816954.1 GI:15911639
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and
AUTHORS Prange, C.K.
TITLE The I.M.A.G.E. Consortium quality control effort: clone
JOURNAL resequencing for verification
COMMENT Unpublished (2001)
Other ESTs: BG058779
Contact: Prange CK
The I.M.A.G.E. Consortium
Lawrence Livermore National Laboratory
Livermore, CA, USA
Email: help@image.llnl.gov
This read has been verified (found to hit its original self in the
correct orientation), as part of the I.M.A.G.E. Consortium quality
more base pairs with a phred quality value of 20 or greater, where
a sliding window of 4 base pairs with a phred quality value of 15
or greater marks the beginning and end of the sequence. For
information on obtaining this clone, please contact
info@image.llnl.gov. effort.
Plate: L1AM9388 row: k column: 7
Seq primer: -21ml3
High quality sequence stop: 682.
Location/Qualifiers
1..682
```

FEATURES

```
Source
1..682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4140798"
```

```
/tissue_type="lymphocyte"
/lab_host="pH10B (phage-resistant)"
/clone_id="Soares_NPBMC"
/note="Organ: blood; Vector: pT73D-Pac; Site 1: NotI;
Site 2: EcoRI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTATCCATCTGAAGTGGGAGCGCGCGGGTGTGTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
is normalized; constructed in the laboratory of M. Bento
Soares (University of Iowa)."
```

ALIGNMENT SCORES:

Pred. No.:	7,44e-119	Length:	682
Score:	1086.50	Matches:	214
Best Local Similarity:	96.89%	Conservative:	4
Best Local Similarity:	95.11%	Mismatches:	6
Query Match:	69.47%	Indels:	2
DB:	12	Gaps:	1

US-10-029-345A-109_COPY_1_302 (1-302) x BI816954 (1-682)

```
QY 49 AsnCysSerLysLeuMetLysArgArgLeuGlnIleAspLysValLeuIleThrGlnLeu 68
Db 673 AACTGTCACAGCTA--TAAAGCAAGTGCAACAGACAAAGTGT-ATTAACAGGCTC 618
QY 69 IleGlnHisSerAlaLysHisLysValAlaPheIleAspCysSerGlnLysValAlaTyr 88
Db 617 ATCCAGGATTCACGCAACATTAAGCTTGACATTTGATGACAGACAGAGGTGTAGTTAC 558
QY 89 AspGlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeuThrValLeuLeu 108
Db 557 GATCAAGAGCTCCCAAAATGTCCTCTCTCTTCAAGATGTTTTCTACCTGATTTCTG 498
QY 109 G1YLeuGlnIleLysSerPheAsnSerValHisLeuAlaGlyGlyPheAlaGluPhe 128
Db 497 GGTAACTGGAGAGAGCTTCACTGTGTCACCTGTCGACGAGGTGGCTTGACAGTTC 438
QY 129 SerArgCysPheProGlyLeuCysGlnGlyLysSerThrLeuValProThrCysIleSer 148
Db 437 TCTCGTGTGTTCCCTGGCTCTGTGAAGAAATCACTAGTCCCTACCTGATTTCT 378
QY 149 GlnProCysLeuProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyrLeu 168
Db 377 CAGCTTGCTTACCTGTGGCCAACTTGGGCCAACCCGAATCTTCCCAATCTTATCTT 318
QY 169 G1YCysGlnArgAspValLeuAsnLysGlnLeuIleGlnIleAsnGlyTleGlyTyrVal 188
Db 317 GGCTGCCAGCGAGATGCTCTCAACAAGAGCTGATGACGAGATGGAGTTGGATTATGTG 258
QY 189 LeuAsnAlaSerTyrThrCysProLysProAspPheIleProGlnSerHisPheLeuArg 208
Db 257 TTAATGTCAGCAATACCTGTGCCAAAGCTGACTTATCCCGAGTCTCATTTCCGCT 198
QY 209 ValProValAsnAspSerPheCysGlnLysIleLeuProThrLeuAspLysSerValAsp 228
Db 197 GTGCTGTGATGACAGCTTTTGTGAAGAAATTTTGGCGGTGTGGCAATATAGTAGAT 138
QY 229 PheIleGlnIleValAlaValAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIle 248
Db 137 TTTCAITGAGAAAGCAAAAGCCTCCAAATGATGATGTTCTTATGACAGTGTAGCTGGAGATC 78
QY 249 SerArgSerAlaThrIleAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAsp 268
Db 77 TCCGCTGCCGACCAATGCTATGCTTACATCATATGAAGATGACATGTCTTATGAT 18
QY 269 GlnAlaTyrArgPhe 273
Db 17 GAACTTACAGATT 3
```

RESULT 5

CF532917 656 bp mRNA linear EST 12-SEP-2003
DEFINITION UI-M-GH0-cgw-n-02-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone
IMAGE:30357529 5', mRNA sequence.
ACCESSION CF532917
VERSION CF532917.1 GI:34584885
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 656)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
FEATURES
source
Location/Qualifiers
1..656
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30357529"
/tissue_type="whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH_BMAP_GH0"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
Site 2: Not I; The library was constructed according
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGACTGAAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN
Alignment Scores:
Pred. No.: 3,17e-118 Length: 656
Score: 1081.00 Matches: 208
Percent Similarity: 98.60% Conservative: 4
Best Local Similarity: 96.74% Mismatches: 2
Query Match: 69.12% Indels: 1
DB: 14 Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x CF532917 (1-656)

QY 16 ValAlaLeuLeuGusUserGlyThrGluValLeuLeuLeaPSeArGPro-PheVa 35
Db 3 GTGGCTCTGGCTGGAAGTGAACGGAABAAAGTGTCTAATTATGATCCGACATTTGT 62
QY 35 lGluTYrAsnThSeerHsilleLeuGUAlaAlaAsnilleAsnCYSeSerlySeuWecly 55
Db 63 GGAATACAAATACGCTCACAATTGGAAAGCATTAATAATCAACGCTCCAAACGATGAA 122
QY 55 bArGArGlueGlnGlnApLybValLeuileThrGluLeuileGlnHsSerAlaLybH 75

Db 123 GCGAGGTGGCAACAGACAAAGATTAATTAACAGAACTTATCAACATCTCGAAAGCA 182
QY 75 bLySValaPleleAPCySeSerGlnlyValVallyTYrAsGlnSeSerGlnApLy 95
Db 183 TAAAGTTGACATTGATGATGCAAGAGGGTGAATTATGATCAAAAGTTCCAAAGTGT 242
QY 95 lAlaSerLeuSeSerAPCyAPhLeuThrValleLeuGlyLyLeuGlyLysSerPh 115
Db 243 TGGTTCTGTGCTGACAGCTGCTTCTCACTGACTTCTGGGTAACCTGAGAGACCTT 302
QY 115 eaenSerValHsleLeuAlaGlyGlyPhealagluPheSeSerArGyAPhProGlyle 135
Db 303 CAACCTGTGCACTGCTGAGGTGGCTTGTGAGTTCTCGTTGTTTCCCTGGCCT 362
QY 135 uCYsGluGlyVSeSerThreValAlProthCyVllSeSerGlnProCyLeuProValAl 155
Db 363 CTGTGAAGGAAGATCCACTCTACTCTCCCTACCTGCAATCTCACCTGCTTACTCTGTGC 422
QY 155 aAenilleGlyProthArGilleuProAenleuTYrleuGlyCYsGlnArGAsPValle 175
Db 423 GAACATTGGCCCACTGAAATTTCTCCAACTCTATCTTGCTGCGAGGAAATGCTT 482
QY 175 uAenilleGlyleuileGlnGlnAnGlylleglyTYrValleuAAnlaSeerTYrThry 195
Db 483 CAACAAGACCTGATGACACAGATGGANTGGCTATGTTAAAGCCAGCAATACCTG 542
QY 195 sProlYsPProAPhelleProGluSerHsPheleuArGValProValaPnApSePh 215
Db 543 TCCAAAGCCGACTTATCACTGAATCTCACTTCCGCGAGTCCGTGTGAATGACAGCTT 602
QY 215 eCYsGluLyVleleuProThreApLybVallePhe 229
Db 603 TTGTGAAGAAATCTACCATGTTGAGCAAGTGTGAGATTTC 645

RESULT 6
CF742387 639 bp mRNA linear EST 10-OCT-2003
DEFINITION UI-M-HB0-cl1-C-20-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
IMAGE:30619363 5', mRNA sequence.
ACCESSION CF742387
VERSION CF742387.1 GI:37638726
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 639)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
FEATURES
source
Location/Qualifiers
1..639
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30619363"
/tissue_type="whole eye"
/dev_stage="embryo 12.5, 13.5, 14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH_BMAP_HB0"

Db 126 CTGGCCCTGTGAGGAAGTCACTCTAGTCCCTACTGATATCTCAGCTTGCTTA 185
Qy 153 ProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyrIleuGlyCysGlnArg 172
Db 186 CCGTGGCGAAGCTGGCCCAACTCGAATCTCCCAATCTCTATCTTGCTGCGACGGA 245
Qy 173 AspValLeuAsnTyrGlyLeuIleGlnIleAsnGlyIleGlyTyrValLeuAsnAlaSer 192
Db 246 GATGCTCTCAACAGAGCATGATGCAACAGATGGATTTGGCTATGTGTTAAATGCCAGC 305
Qy 193 TyrThrCysProIysProAspPheIleProGlySerHisIleLeuArgValProValAsn 212
Db 306 AATACCTGCTCAAGAGCTGACTTCACTGATCTGATCTCTCCGAGAGCTGTGAAT 365
Qy 213 AspSerPheCysGlyLeuIleuProThrLeuAspIysSerValAspPheIleGlyLeu 232
Db 366 GAGAGCTTTGTGAGAAATCTCAACATGCTGAGCAAGCTGTGAGATTCATTGAGAAAA 425
Qy 233 AlaIleAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAla 252
Db 426 GCAAAAGCTCCCAATGCTGCTGCTTATCTCATCTGCTTACCTGGATCTCTCCGCC 485
Qy 253 ThrIleAlaIleAlaTyrIleMetIysArgMetAspMetSerLeuAspGlnAlaTyrArg 272
Db 486 ACTATGCTATGTGCTCATCATGAGAGATGACATGCTCTAGATGAGGCTTACAGA 545
Qy 273 PheValIleGlyLeuIysArgProThrIleSerProAsnPheAsnIleuGlyIleuLeu 292
Db 546 TTGTGTGAAGAAAAAAGAACTACTATATCTCCGAATTTTATGAGGCCAATCATG 605
Qy 293 AspTyrGlyLeuIysIleLeuAsnGlnThr 302
Db 606 GACTATGAGAGAGATTATTAATACCAACT 635

RESULT 8
BF472046 650 bp mRNA linear EST-04-DEC-2000
LOCUS UI-M-BH3-awu-d-10-0-UI.r1 NIH BMAP_M_S4 Mus musculus cDNA clone
DEFINITION UI-M-BH3-awu-d-10-0-UI 5', mRNA sequence.
ACCESSION BF472046 GI:11541229
VERSION BF472046
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Bernaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MEST@mail.nih.gov
CDNA library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Berto Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seq primer: M13 Reverse.
FEATURES
Location/Qualifiers
1..650
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

/clone="UI-M-BH3-awu-d-10-0-UI"
/dev stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH_BMAP_M_S4"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH BMAP_M_S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M_S4,
NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
(NIH_BMAP_M_S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH_BMAP_M_S4 library. This procedure has been previously
described (Bernaldo, Lennon and Soares, Genome Research
6:791-806, 1996)"

ORIGIN
Alignment Scores:
Pred. No.: 2,86e-109 Length: 650
Score: 1006.00 Matches: 205
Percent Similarity: 97.21% Conservative: 4
Best Local Similarity: 95.35% Mismatches: 6
Query Match: 64.32% Indels: 3
DB: 10 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x BF472046 (1-650)
Qy 5 MetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlnSerGlyThrGlu 24
Db 3 ATGATGGAGAACTCAAAATTTGTTACTGAGAGCTTGCTGCTGCAAGATGGAACGGA 62
Qy 25 LysValLeuLeuIleAspSerArgProPheValGlyTyrAsnThrSerHisIleGlu 44
Db 63 AAATG-CTGCTAATGATATGCGCAGCATTTGTGGAATACAAATGCTCTCATTTGGAA 121
Qy 45 AlaIleAsnIleAsnCysSerIysLeuMetIysArgArgLeuGlnIleAspIysValLeu 64
Db 122 GCCATTAATATCAACTGCTCCAAAGTGAAGCCGAGAGGTGCAACAGCAAAAGATTAA 181
Qy 65 IleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAspCysSerGlnLys 84
Db 182 ATTACGAACCTTATCCAACTTTGCAAGACATPAGCTTGACATTGACATTCAGAG 241
Qy 85 ValValAlaTyrAspGlnSerSerGlnAspValAlaSerLeuSerAspCysPheLeu 104
Db 242 GTGTGATTATGATCAAGATTCCAAAGATGTTGTTCTGTGCGTCAAGCTGCTTCTC 301
Qy 105 ThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeuLeuAlaGly 124
Db 302 ACTGTCCTTGGGTAACTGAGAGACCTTCACCTGTGCACTGCTGAGAGTGGC 361
Qy 125 PheAlaGluPheSerArgCysPheProGlyLeuCysGlnGlyLysSerThrIleuValPro 144
Db 362 TTTCGTGAGTTCTCTGTTGTTCTCTGCGCTCTGGAAGAAA-TCCACTCTAGTCCCT 420
Qy 145 ThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThrArgIleuPro 164
Db 421 ACTGTGATATCTCAGCTTGCTTACTGTTGGAACATTTGGGCCAATTCGAATTTCTCC 480

OY	165	AshenuryrtlenylyCysglnagapValIleuanilysguleenileginginaemgly	184
Db	481	AATCTCATCTTGCGTCCAGGACGATGTCCCAACAAGCACCGAATGAACAGATGGG	540
OY	185	IleglYrYalIeuanasaserTYrrThrcysProlaspProapPhellelepoguser	204
Db	541	ATTGGCTATGTGTAAATGCAGCAATACTGTCTCCAAGCCTGTACTTCACCTGATCT	600
OY	205	HiephelenurgValIprovalasnspserPhecysglutylsile	219
Db	601	CACITCTCGGAGTGTCTGTGAATGACAC-TTTTGTGTAGAATAATC	644
RESULT 9			
LOCUS	BC038231		
DEFINITION	2207 bp	mRNA	linear HTC 01-OCT-2002
ACCESSION	BC038231		
VERSION	EC038231.1		
KEYWORDS	BC038231.1	GI:23398534	
SOURCE	HTC.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
STRASBERG,R.	1 (bases 1 to 2207)		
Submitted	(30-SEP-2002)	National Institutes of Health	Mammalian

REMARK COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk

cerna Library Preparation: The Life Technologies, Inc.
 CNLA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC)
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nhgri.nih.gov
 Akhter, N., Ayele, K., Beckertom-Sternberg, S. M., Benjamin, B.,
 Blakesley, R. W., Bouffard, G. G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N. L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, B., Kwong, P., Latic, P., Legaspi, R.,
 Maduro, Q. L., Mastello, C., Mackeri, B., Mastrian, S. D., McCloskey, J. C.,
 McDowell, J., Pearson, R., Stathitop, S., Thomas, P. J., Touchman, J. W.,
 Tsugeon, C., Vogt, J. L., Walker, M. A., Wetherby, K. D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E. D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRKX Plate: 79 Row: a Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4758211
This clone has the following problem: frame shifted.
Location: 903612

```

source
1. 2207
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:5547764"
   /tissue_type="Skin, melanocyte"
   /clone_id="NH_MGC_72"
   /lab_host="DH10B"
   /note="Vector: pCMV-Sport6"

```

ORIGIN	1
Alignment Scores:	
Pred. No.:	2,31e-106
Score:	989.00

Percent Similarity:	81.40%	Conservative:	55
Best Local Similarity:	63.12%	Mismatches:	54
Query Match:	63.24%	Indels:	2
DB:	11	Gaps:	2

US-10-029-345A-109_COPY_1_302 (1-302) x BC038231 (1-2207)

OY	1	MetAlaHisgIuMetIleGlyThrGlnIleVal---ThrGlnArgLeuValAlaLeuLeu	19
Db	132	ATGGCTGGGAGACCGGCTCCCGAGAAAGTGTATGGATGTCCAAGAGCTGGCCAGCTCTGTG	191
OY	20	GluSerGlyThrGlnValLeuLeuIleAspSerArgProPheValGluTyraThr	39
Db	192	CGAGGCGCGCTGGGGGCGGTGTGCATCGACAGCGCTCTTCGTGGAGTACACAGC	251
OY	40	SerHisIleLeuGlnAlaIleAsnIleAsnCySerIleuMetIleuValArgArgLeuGln	59
Db	252	TGGCATGTGCTCAAGCTCTGTCAACATCTGTCTCTCCAAAGCTGGTGAAGCGGCGGTGCAG	311
OY	60	GlnAspIleValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisIleValAspIle	79
Db	312	CAGGGAGAGGTGAACATTTGGGAGCTCATTCACCGGCTGCACGACAGCGAGTGAAGGCT	371
OY	80	AspCySerGlnIleValValIleTyraSpGlnSerSerGlnAspValAlaSerIleuSer	99
Db	372	ACGAGACCAACAGACCGTGGTGGTCTATGACAAAGACAGCGGGAGCGCGCTGGGC	431
OY	100	SerAspCyPheLeuThrValLeuLeuGlyIleGlnIleuValSerPheAsnSerValHis	119
Db	432	CGAGACAGCTTCTCTCCATCTCTGCTGACACAGCTGAGCGGCTTTCACAGCGGTGGCC	491
OY	120	LeuLeuAlaGlyIlePheAlaGluPheSerArgCyPheProGlyIleuCySGlnGlyIle	139
Db	492	ATCTCATCGGGGGCTTGGCCACCTTCTCTCTGTTTCCGGGCTCTGCAGAGGCAAG	551
OY	140	---SerThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGly	158
Db	552	CCTGGCTCCCTTCAACCATGAGCGCTCTCCAGCCCTGCTGCTGCTGCCAGCGTGGGC	611
OY	159	ProThrArgIleLeuProAsnLeuTyrlleuGlyCysGlnArgAspValLeuAsnIleGlu	178
Db	612	CTGACCCGGATCTGTGCTCACCTTCACTGGGCTGCGAGAAAGAGCTCTTAACAAAGAT	671
OY	179	IleuIleGlnGlnAsnGlyIleGlyTyraIleuAsnAlaSerTyThrCysProIleAsp	198
Db	672	CTGATGAGCGCAAAATGATTAAGCTACCTCTCAACGCGACGAACTCTGCCCAAGCT	731
OY	199	AspPheIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIle	218
Db	732	GACTTCATCTGCGAGAGCGCGCTTCAATGGGGTCCCCATCAACGACACTGTGAAAAA	791
OY	219	IleLeuProTyrLeuAspIleSerValAspPheIleGluIleValAlaIleValSerAsnGly	238
Db	792	CTGGCGCCTGGCTGAGCAAGTCCATGAGTTTCATGATTAAGCCAGAGCTTCCAGCTGC	851
OY	239	CysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTy	258
Db	852	CAGGTCATGCTCACTGCTGGCGGAGATCTCCCGCTCTGCCACATTCGCATCGCTAC	911
OY	259	IleMetIleArgMetCAspMetSerLeuAspGlnAlaTyraArgPheValIleGluIleArg	278
Db	912	ATCATGAAGACCAAGGCGATGTCTCTCCAGACAGCCCTTACAGGTTCGTGAAGACAGCGC	971
OY	279	ProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrglyIleValIle	298
Db	972	CCGTTCATCTGCGCCAACTTCAACTTCTGGGCGAGCTGTGGAGTACGAGCGAGCTGC	1031
OY	299	Ile	
Db	1032	AAAG	1034

LOCUS	924 bp	mrna	linear	EST 12-NOV-2002
DEFINITION	CA474739	Agencourt 10700573	NCI CGAP Zfid1	Danio rerio cdna clone
ACCESSION	CA474739	IMAGE:6757004	5'	mrna sequence.
VERSION	CA474739.1	GI:24931091		
KEYWORDS	EST.			
SOURCE	Danio rerio (zebrafish)			
ORGANISM	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.			
REFERENCE	1 (bases 1 to 924)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-rc@mail.nih.gov Tissue Procurement: Leonard I. Zon, M.D. cDNA Library Preparation: Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLI at: http://image.llnl.gov Plate: LLAMI4309 row: j column: 11 High quality sequence start: 24 High quality sequence stop: 708.			

FEATURES
SOURCE

```

/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon,7955"
/clone="IMAGE:6397004"
/lab_host="DH10B (TI-resistant)"
/clone_1fb="NCI CGAP Zf1d1"
/notes="Organ: kidney; PCMV-SPORT6.1; Site 1:
ECOR3; Site 2: NciI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.8 kb. Constructed by J.
Wang (Research Genetics, Invitrogen Corp) from tissue
donated by L. Zon (Harvard University). Note: this is a
NCI_CGAP Library."

```

ORIGIN

Alignment Scores:	
Pred. No.:	9.89e-106
Score:	978.50
Percent Similarity:	86.08%
Best Local Similarity:	74.36%
Query Match:	62.56%
DB:	14
Length:	92.92
Matches:	20.50
Conservative:	32
Mismatches:	33
Indels:	6
Gaps:	1

US-10-029-345A-109_COPY_1_302 (1-302) X CA474739 (1-924)

Oy	113	lysSerpheAnsSerValIHIsleLeuValAGIyPheAlaGlyPheAlaGlyPheSerArgCyAphe	132
Db	415	AAGAGTTTCCCATCAGTCCACCTCTCTCAGGGGGTTTCTCGAGATTCTCAGGGCTTTC	474
Oy	133	ProGlyLeuCyGsgIuglyIyysSerThrLeuValProThrCySIIesGlnProCyLeu	152
Db	475	CCGGGCTGTGTAGAGGGGAAGTCTGCGCTGGTCCGTCTGTGTGTGAGAGTGTGTGTG	534
Oy	153	ProValAlaAsnIIeglyProThrArgIIeLeuProAsnLeuTyrLeuGlyCyGsgIIArg	172
Db	535	TCCGTCTCAGACGCGCGGGGCCACCGCGGATCTCTGCACACCTGACCTCGGGCTGCACGG	594
Oy	173	AspValLeuAsnIyysgIleuIIegIInGlnAsnGlyIIeglyTyrValLeuAsnAlaSer	192
Db	595	GACCTGTCTCAACAGAGCTGATGACACAGACAGACATCCGCTTCTGTGTAGACCGACG	654
Oy	193	TyrThrCySProluysProAspPheIIeProGlySerIIAspHeuArgValProValAsn	212
Db	655	AACCTCTGCCCAAGCCAGACTTCATCCCGACACACACTTCTCGCGGGTGCCTGTAAAC	714
Oy	213	AspSerPheCyGsgIuysIIeLeuProTrpLeuAspIyysSerValAspPheIIegIuys	232
Db	715	GACAGCTTTCGCCGAGAGATCTCTGCGTGGCTGCACGCTCCGCGAGTTCATGAGAAA	774
Oy	233	AlaIyysAlaSerhengIyCySvalLeuValIHIsCyLeuAlaGlyIIeserArgSerAla	252
Db	775	GCCAAAGCCAGTAACGCCAGAGTTCTGTTCATCTGTCTGGCGGGAAATCTCCCGTCCGCC	834
Oy	253	ThrIIeAlaIIeAlaTyrIIeMet-LysArgMetAspMetSerIeu-AspGluAlaTyr-	271
Db	835	ACATC-GCCATCGCTCACTCATTAAGAAGATGACATGACACTGGGATTAAGACGTACC	893
Oy	272	ArgPhe-ValIyysGluysArgProThrIIe	281
Db	894	AGGTTCCGGAAGGAGAACGCGCGACATC	924

RESULT	11
BQ770036	862 bp mRNA linear EST 26-JUL-2002
LOCUS	
DEFINITION	UI-M-F10-b-ytc-o-24-0-UT.r1 NIH BMP_F10 Mus musculus cDNA clone IMAGE:5702255 5', mRNA sequence.
DESCRIPTION	

ACCESSION	BQ770036	
VERSION	BQ770036.1	GI:21978510
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

REFERENCE	1 (bases 1 to 862)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strauberg, Ph.D.

Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: PYX-5, GCGTGGGCTGGAGTTGGTGATG

FEATURES

```
source
1. .862
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="C57BL/6"
  /db_xref="taxon:10090"
  /clone="IMAGE:5702255"
  /tissue_type="whole brain"
  /dev_stage="embryo 12.5dpc"
```


|||||
Db 414 GTTCACCTGCTTGCAGGTGGGTTTGGCTGAGTCTCTCGTTGTTCCCTGGCCTCTGAA 355
|||
Qy 138 GtlyvserthrleuvalprothrCysIleserGlnProCysleuProValAlaalanile 157
Db 354 GGAAAAATCCACTTACCTCTACCTGATTTCTCAGCTGCTTACCTGTTGGCCAACTT 295
|||
Qy 158 GtlyProthrArgIleleuProanleuTyrlleuGtlyCysGlnArgApValleuanly 177
Db 294 GGCCCAACCCGAAATCTTCCCAATCTTATCTTGGCTGCCAGAGATGCTTCAACAG 235
|||
Qy 178 GtlyleuIleGlnGlnanGlyIleGtlyTyrlValleuanAlaSerTyrrThrCysProlys 197
Db 234 GACCTATACAGCAGAAATGGGATTTGGTATGTTAAATCCAGCTATACCTCTCCNAAG 175
|||
Qy 198 ProAspPheIleProGluSerHisPheleuArgValProValanAspSerPheCysGlu 217
Db 174 CCGACCTTATCCCGAGTCTCATTTCTGCTGCTGCTGCTGATGACAGCTTTTGTGAG 115
|||
Qy 218 TyrlleuProthrleuAspIlySerValAspPheIleGtlyValleuanAlaSerAn 237
Db 114 AAAATTTTGGCGGTGGACAA-TCAGTAGATTTCATTAGAAAGCAAGCCTC--CAT 58
|||
Qy 238 GtlyCysValleuValHisCysleuAlaGtlyIleSerArgSerAlaThr 253
Db 57 GGATGCTTCTACTGACCTGTTTACCTGGCATCTCCCGCTCCGCCACC 10
|||
RESULT 13
AM847426 595 bp mRNA linear EST 19-MAY-2000
LOCUS R01-CT0206-270999-021-g06 CT0206 Homo sapiens cDNA, mRNA sequence.
DEFINITION AM847426
ACCESSION AM847426 GI:7942943
VERSION AM847426.1 GI:7942943
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 595)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Negal,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PubMed 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPSP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?lcl=RC1-CT0206-270
999-021-g06&lcl=1999-09-27&lcl=1)
Seq primer: puc 18 forward
High quality sequence stop: 595.
Location/Qualifiers
1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0206"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived

|||||
Db 414 GTTCACCTGCTTGCAGGTGGGTTTGGCTGAGTCTCTCGTTGTTCCCTGGCCTCTGAA 355
|||
Qy 138 GtlyvserthrleuvalprothrCysIleserGlnProCysleuProValAlaalanile 157
Db 354 GGAAAAATCCACTTACCTCTACCTGATTTCTCAGCTGCTTACCTGTTGGCCAACTT 295
|||
Qy 158 GtlyProthrArgIleleuProanleuTyrlleuGtlyCysGlnArgApValleuanly 177
Db 294 GGCCCAACCCGAAATCTTCCCAATCTTATCTTGGCTGCCAGAGATGCTTCAACAG 235
|||
Qy 178 GtlyleuIleGlnGlnanGlyIleGtlyTyrlValleuanAlaSerTyrrThrCysProlys 197
Db 234 GACCTATACAGCAGAAATGGGATTTGGTATGTTAAATCCAGCTATACCTCTCCNAAG 175
|||
Qy 198 ProAspPheIleProGluSerHisPheleuArgValProValanAspSerPheCysGlu 217
Db 174 CCGACCTTATCCCGAGTCTCATTTCTGCTGCTGCTGCTGATGACAGCTTTTGTGAG 115
|||
Qy 218 TyrlleuProthrleuAspIlySerValAspPheIleGtlyValleuanAlaSerAn 237
Db 114 AAAATTTTGGCGGTGGACAA-TCAGTAGATTTCATTAGAAAGCAAGCCTC--CAT 58
|||
Qy 238 GtlyCysValleuValHisCysleuAlaGtlyIleSerArgSerAlaThr 253
Db 57 GGATGCTTCTACTGACCTGTTTACCTGGCATCTCCCGCTCCGCCACC 10
|||
RESULT 14
AK035652 3325 bp mRNA linear HTC 19-SEP-2003
LOCUS AK035652
DEFINITION Mus musculus adult male urinary bladder cDNA, RIKEN full-length
enriched library, clone:9530081F05 product:RUNKATBD MAPK
PHOSPHATASE 7 homolog [Homo sapiens], full insert sequence.
ACCESSION AK035652
VERSION AK035652.1 GI:26330815
KEYWORDS HTC; GAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

JOURNAL MEDLINE 99279253
REFERENCE 10349636
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL MEDLINE 20499374
PUBMED 11042159
AUTHORS 3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitesunai,T., Tashiro,H., Itoh,M., Suni,N., Ishi,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanuki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL MEDLINE 20530913
PUBMED 11076861
AUTHORS 4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL MEDLINE 11076861
PUBMED 11076861
AUTHORS 5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL MEDLINE 11076861
PUBMED 11076861
AUTHORS 6
(bases 1 to 3325)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Hori,P., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL MEDLINE 11076861
PUBMED 11076861
AUTHORS 1
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
FEATURES SOURCE
Location/Qualifiers
1..3325
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="FANTOM,DB:9530081F05"
/db_xref="MG1:2399973"
/db_xref="taxon:10090"
/clone="9530081F05"

CDs
/sex="male"
/tissue_type="urinary bladder"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
/note="unnamed protein product; TRUNCATED MAPK PHOSPHATASE 7 homolog [Homo sapiens] (SPR1096082, evidence: PASTY, 96.7%id, 86%length, match=369)
putative"
/codon_start=1
/protein_id="BAC29138.1"
/db_xref="GI:26330816"
/translation="MAHEMIGTGVITVESLVALLSGETKVLIDSPPEVNTSHILE
AIVNLSKIMKRLQODKVLITELIQHSAAHKVIDCONKVVVDSSQDVGSLSDDC
FLVNLGKLSRFSNVHLLAGSPFRCPGCEKSTLVPICISQPCIPVANIIGPT
RLPNLVIGGQRLVKNVGSKHLISLCTLSVASKAAVYHLSNII"

ORIGIN
Alignment Scores:
Pred. No.: 2,34e-94 Length: 3325
Score: 891.00 Matches: 173
Percent Similarity: 98.87% Conservative: 2
Best Local Similarity: 97.74% Mismatches: 2
Query Match: 56.97% Indels: 0
Gaps: 11
US-10-029-345a-109_copy_1_302 (1-302) x AK035652 (1-3325)
QY 1 MetAlHisgluWetlleglyThrGlnlleValThGluArgleuValAlaleuLeuGlu 20
Db 460 ATGCCCATAGATGATGGAACTCAATTTGTACTAGAGAGCTGGGCTGCTGCGAA 519
QY 21 SerGlyThrGluysValleuLeuileAspSerArgProPheValGluTyAsnThrSer 40
Db 520 AGTGAACGAAAGAGCTGCTTAATGATGACGACCATTTGTGCAATCACTACTCT 579
QY 41 HislleuGluAlaleuAsnleAsnCySerlylsleuWetlysaArgleuGlnGln 60
Db 580 CACATTTGGAAGCATATATATCACTGCTCCAACTGATGAAGCAAGGTTGCAACAG 639
QY 61 AsplysValleuileThrGluileuileGlnHisSerAlaGlnHislylsValAsp 80
Db 640 GACAAATATTAATTCAGAACTTATCCACATTCGCAAGCAATGAGTGAACATTGAC 699
QY 81 CyserGlnlysaValValAlaTyAspGlnSerSerGlnAspValAlaSerLeuSer 100
Db 700 TGCAATGAGAGGTGATGATTTATGATCAAGTTCCCAAGATGTGTTCTCTGCTCA 759
QY 101 AspCyPheleuThrValleuLeuGluGlylsleuGluysSerPheAsnSerValHisleu 120
Db 760 GATGCTTTCTCACTGTAATTCGGTAACTGAGAGGAGCTTCAACTGTCCTCACTG 819
QY 121 LeuAlaGlylyPheAlaGluPheSerArgCyPheProGlyleuCyseGluGlylySer 140
Db 820 CTTCGAGGTGGCTTGTGTAATTCCTGTTGTTCCCTGCGCTTGTGAAGAAAGTCC 879
QY 141 ThrleuValProThrCysIleSerGlnProCysleuProValAlaAsnIleGlyProThr 160
Db 880 ACTCTAATCCCTCACTGCAATATCTCACTGCTTCTTACTTTCGAAATATGGGCCACT 939
QY 161 ArgIleleuProAsnleuTyTyLeuGlyCyseGlnArgAspValleuAsnly 177
Db 940 CGAATTTTCCCAATCTATCTTGGTGCAGAGATGTCTCAACAAG 990
RESULT 15
BU946569/c 792 bp mRNA linear EST 17-OCT-2003
LOCUS 703769910J1 RALUTYX101 Rattus norvegicus cDNA, mRNA sequence.
DEFINITION BU946569
ACCESSION BU946569
VERSION BU946569.1 GI:37701886
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

AUTHORS

1 (bases 1 to 792)
Fu, G.K. and Stuve, L.L.
An improved method for the construction of cDNA libraries for highly efficient DNA sequencing from the 3' end of expressed genes

JOURNAL

Unpublished (2003)

COMMENT

Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.

FEATURES

source
location/Qualifiers
1..792

/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/issue_type="Lung"
/clone_lib="RALUTX101"
/note="Rat, Lung, Clotifibrate, 250 mg/kg, 12hr-day 28 sacrifice, SD, M/F, Pool"

ORIGIN

Alignment Scores:

Pred. No.:	4.09e-93	Length:	792
Score:	872.00	Matches:	178
Percent Similarity:	93.78%	Conservative:	3
Best Local Similarity:	92.23%	Mismatches:	9
Query Match:	55.75%	Indels:	4
DB:	13	Gaps:	0

US-10-029-345a-109_COPY_1_302 (1-302) x BU946569 (1-792)

```
QY      82  SerGlnIysValValValIYrAspGlnSerSerGlnAspValAla-SerLeuSerSerAs 101
      |||
Db      779  AGTCAGAGAGGTGGTAGTATGATGATCAGATTCACAGATGGTTCGGTCCGTCCAGACC 720
QY      101  pCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSer-ValHisLeuL 121
      |||
Db      719  GGCCTTCTTCACGTGCTTCGTGGGTAGCTGAGAGAGCTTCACACTGCTCACCCTGC 660
QY      121  euAlaGlyIysPheAla-GluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
      |||
Db      659  TTGCAGGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
QY      141  ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
      |||
Db      599  ACTCTAGTCCCTACCTGATATCTGACGCTTGGCTTACCTTGGCAACACTGGGCGCACT 540
QY      161  ArgIleLeuProAsnLeuIYrLeuGlyCysGlnArgAspValLeuAsnIYrGluLeuIle 180
      |||
Db      539  CGAATCTCTCCCATCTCTATCTCGGCTCCAGAGATGCTCCCAAGAGAGCTGATG 480
QY      181  GlnGlnAsnGlyIleGlyIYrValLeuAsnAlaSerIYrThrCysProIYrProAspPhe 200
      |||
Db      479  CAACAGAACGGAGATGGCTATGCTAAATGCCAGCAACCTGTCACAAAGCCTGACTTC 420
QY      201  IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIYrIleLeu 220
      |||
Db      419  ATACCGGAATCTCACTTCCGCGAGTCCCTGTGATGACAGCTTTTGTGAGAAATCTTG 360
QY      221  ProTrpLeuAspIYrSerValAspPheIleGluIYrAlaLysAlaSerAsnGlyCysVal 240
      |||
Db      359  CC-TGGTTGGACAAAGTGTGGATTTCATTCAGAAAGCAAAAGCTCCAAAGCGCTGTG 301
QY      241  LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIYrIleMet 260
      |||
Db      300  CTGATTCACGTGCTGGGATCTCCGCTCCGCAACCAATGCTATTGCTATATCATG 241
QY      261  LysArgMetAspMetSerLeuAspGluAlaIYr 271
      |||
Db      240  AAGAGATGACATGCTCTGTGATGAGGCTTAC 208
```

Search completed: June 21, 2004, 23:33:32
Job time : 2267.97 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 19:56:46 ; Search time 334.168 Seconds
(without alignments)
3839.259 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMIGTIVTERLVALLE.....PNNFGLGDLDEKIKIKQT 302

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10029345/rnat 21062004.164145.29020/app_query.fasta_1.1294
-DB=Genseq 29345 -QFMT=fasta -SUPLX=eng -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human0.cdi
-LIST=45 -DOCALL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345 -CCGN 1.1.687 -runat 21062004.164145.29020 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04:.*
1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001a:.*
5: geneseqn2001b:.*
6: geneseqn2002a:.*
7: geneseqn2003a:.*
8: geneseqn2003b:.*
9: geneseqn2003c:.*
10: geneseqn2004a:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1564	100.0	5450	6	ACC60559 Polynucle
2	1564	100.0	5450	6	ACC60572 Polynucle
3	1552	99.2	2118	4	AA30479 Human pro
4	1552	99.2	2732	4	AAD09492 Human SGP
5	1552	99.2	3059	6	AAS15768 Human sequ
6	1552	99.2	3104	6	ABNS9704 Novel hum
7	1552	99.2	3496	6	ABK47596 CDNA enco
8	1552	99.2	3544	5	AAS14639 Human CDN

9	1552	99.2	3544	6	ABK49402 CDNA enco
10	1552	99.2	3766	6	ABK14474 Human pro
11	1552	99.2	4790	6	ABN83966 Human gen
12	1552	99.2	5145	5	ABV20833 Human pro
13	1552	99.2	5145	5	ABV21080 Human pro
14	1552	99.2	5145	5	ABV26680 Human pro
15	1552	99.2	5145	5	ABV20978 Human pro
16	1552	99.2	5145	5	ABV21092 Human pro
17	1552	99.2	5145	5	ABV21312 Human pro
18	1552	99.2	5145	5	ABV21316 Human pro
19	1552	99.2	5145	5	ABV26826 Human pro
20	1552	99.2	5145	5	ABV27131 Human pro
21	1552	99.2	5145	5	ABV26923 Human pro
22	1552	99.2	5145	5	ABV27135 Human pro
23	1552	99.2	5145	5	ABV28657 Human pro
24	1552	99.2	5145	5	ABV22827 Human pro
25	1552	99.2	5145	5	ABV26934 Human pro
26	1552	99.2	5145	5	ABV23105 Human cod
27	1544	98.7	2966	4	AAH99685 Human pro
28	1544	98.7	5111	6	ACC60521 Polynucle
29	1482	95.4	2756	6	ACC60560 Polynucle
30	1246	79.7	749	4	AAH06539 Human CDN
31	1194.5	76.4	3332	6	ABK48378 CDNA enco
32	988	63.2	2377	7	ACA64956 Human pro
33	988	63.2	2377	7	ABX10760 Human du
34	985	63.0	2453	2	AAH86758 CDNA of t
35	985	63.0	2453	2	AAH41236 Murine ne
36	943	60.3	2476	4	AAH31013 Human d
37	934.5	59.8	2415	2	AAH86757 CDNA of t
38	742.5	47.5	1171	5	AAH72088 DNA encod
39	724	46.3	418	5	ABV11252 Human pro
40	724	46.3	461	5	ABV40981 Human pro
41	724	46.3	461	5	ABV32398 Human pro
42	724	46.3	461	5	ABV41327 Human pro
43	715.5	45.7	1917	9	ADH07547 Novel cod
44	698	44.6	427	5	ABV10887 Human pro
45	635	40.6	419	5	ABV32042 Human pro

ALIGNMENTS

RESULT 1	ACC60559	ABK49402
ID	ACC60559 standard; cDNA; 5450 BP.	ABK14474 Human pro
AC	ACC60559;	ABN83966 Human gen
XX		ABV20833 Human pro
DT	19-JUN-2003 (first entry)	ABV21080 Human pro
XX		ABV26680 Human pro
DE	Polynucleotide relating to the invention SEQ ID NO: 108.	ABV20978 Human pro
XX		ABV21092 Human pro
KW	Gene; AB; antiproliferative; hepatocytic; nephrocytic; antiarthritic;	ABV21312 Human pro
XX	antiproliferative; cardiac; cystic; gene therapy; liver disease;	ABV21316 Human pro
KW	proliferative disorder; renal failure; cardiovascular disorder;	ABV26826 Human pro
XX	immunological disorder; arthritis; psoriasis; congenital heart defect;	ABV27131 Human pro
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	ABV26923 Human pro
OS	Homo sapiens.	ABV27135 Human pro
XX		ABV28657 Human pro
PN	WO200257460-A2.	ABV22827 Human pro
PD	25-JUL-2002.	ABV26934 Human pro
XX		ABV23105 Human cod
PF	20-DEC-2001; 2001WO-US050459.	AAH99685 Human pro
XX		ACC60521 Polynucle
PR	20-DEC-2000; 2000US-0256868P.	ACC60560 Polynucle
PR	30-MAR-2001; 2001US-0280186P.	AAH06539 Human CDN
PR	01-MAY-2001; 2001US-0287735P.	ABK48378 CDNA enco
PR	05-JUN-2001; 2001US-0295848P.	ACA64956 Human pro
XX	25-JUN-2001; 2001US-0300465P.	ABX10760 Human du
XX		AAH86758 CDNA of t
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.	AAH41236 Murine ne
XX		AAH31013 Human d

PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
PI Krystek S, Mcatee P, Suchard S, Banas D,
XX WPI: 2002-599721/64.
DR P-PSDB; ABR52381.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
XX
PS Claim 1; Fig 13; 801pp; English.
XX
XX The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
SQ
SQ Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.36e-176 Length: 5450
Score: 1564.00 Matches: 302
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-029-345a-109_COPY_1_302 (1-302) x ACC60559 (1-5450)
QY 1 MetAlHISGLuMetIleGlyThrGlnIleValIThrGluArgLeuValAlaLeuLeuGlu 20
Db 538 ATGGCCCATGAGTATGATGGAATCAATATTTACTGAGAGGTGTGGCTCTGCTGAA 597
QY 21 SerGlyThrGluValLeuLeuIleApsSerArgPropheValGluTyrAsnThrSer 40
Db 598 AGTGAACCGAAAGAGTGTCTGCTAATTGATAGCCGCGCATTTGGCAATCAATACATCC 657
QY 41 HISIleLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
Db 658 CACATTTTGGAGCCATTATATTCACAGCTCCCAAGCTTATGAAAGCCAAAGCTCCAA 717
QY 61 AapLysValLeuIleThrGluLeuIleGlnIleSerAlaIleValIleValIleValIle 80
Db 718 GACAAAGGTGTTAATTAACAGAGCTCATCCAGCATTCAGGAACAATAAGTTGACATTGAT 777
QY 81 CysSerGlnIleValValValValTyrAspGlnSerSerGlnIleValIleValIleVal 100
Db 778 TGCAGTCAAGAGGTGTGATGATCAAGATCCCAAGATGTGCTCTCTCTCTCA 837
QY 101 AapCysPheLeuThrValLeuLeuGlyIleLeuGluIleValIleValIleValIleVal 120
Db 838 GACTGTTTTCTCACGTGATCTCTGGGTAACTGAGAGAGAGCTTCAACTGTTCACTG 897
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyIleSer 140
Db 898 CTTCAGAGGTGGGTGTGCTGAGTGTCTCTGCTGTTCCTCTGCTGAGGAAATATCC 957
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 958 ACTCTAGTCCCTACTGATCTTCTCAGCTTGTCTACCTGTTCTGTTCCCAACATTTGGGCCAAC 1017
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGluLeuIle 180

Db 1018 GCAATTTTCCCAATCTTATCTTGCTGCCAGGAGATGCTTCAACAGGAGCTGATA 1077
QY 181 GlnGlnAsnGlyIleGlyTyrValIleuAsnAlaSerTyrThrCysProLysProAspPhe 200
Db 1078 CAGCAGAAATGGAGATGGTATGATGTTAAATGCCAGTATATCCTGTCCAAAGCCGACTTT 1137
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIleValLeu 220
Db 1138 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGATATACAGCTTTGTGTGAGAAATTTTG 1197
QY 221 ProTrpLeuAapLysSerValIleAspPheIleGluIleValIleValIleValIleVal 240
Db 1198 CCGTGTGTGGACAATACAGTATGATTTTATGAGAAAGCAAAACCTTCAATGATGTGTT 1257
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db 1258 CTAGTGACTGTGTAGGTGGATCTCCCGCTCCGCAACATCGCTATCCCTACATCATG 1317
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluIleValIleVal 280
Db 1318 AAGAGATGGACATGCTTTAGATGAGCTTACAGATTTGTGAAAGAAAGAAAGAACTACT 1377
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuAspTyrGluIleLysIleLysAsn 300
Db 1378 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGACTGATGAGAAAGAAATTAAGAAC 1437
QY 301 GlnThr 302
Db 1438 CAGACT 1443
RESULT 2
ACC60572
ID ACC60572 standard; cDNA; 5450 BP.
XX
XX ACC60572;
XX
XX
XX 19-JUN-2003 (first entry)
XX
XX
XX Polynucleotide relating to the invention SEQ ID NO: 147.
XX
XX Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic;
XX antipsoriatic; cardiant; cytostatic; gene therapy; liver disease;
XX proliferative disorder; renal failure; cardiovascular disorder;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
XX Homo sapiens.
XX
XX WO200257460-A2.
XX
XX 25-JUL-2002.
XX
XX 20-DEC-2001; 2001WO-US050459.
XX
XX 20-DEC-2000; 2000US-0256868P.
XX 30-MAR-2001; 2001US-0280186P.
XX 01-MAY-2001; 2001US-0287735P.
XX 05-JUN-2001; 2001US-0285848P.
XX 25-JUN-2001; 2001US-0300465P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX
XX Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
PI Krystek S, Mcatee P, Suchard S, Banas D,
XX WPI: 2002-599721/64.
DR P-PSDB; ABR52407.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.

XX Example 7, Fig 19, 801pp; English.

PS The invention relates to a novel isolated nucleic acid comprising a
 XX polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotide fully defined in the specification. The polynucleotide of
 CC the invention has antiproliferative, hepatotropic, nephrotropic, the
 CC antitumor, antiproliferative, cardiant, and cytostatic activity. The
 CC polynucleotide may have a use in gene therapy. A polynucleotide or
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferating disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention

XX Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Query Match:	Indels:	Gaps:
1.36e-176	5450	1564.00	302	0	100.00%	0	100.00%	0	0

US-10-029-345A-109_COPY_1_302 (1-302) x ACC60572 (1-5450)

QY 1 MetAlhIeGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 Db 538 ATGGCCCATGAGATGATGGAACTGAAATGTTTACGAGAGGTGGCTGCGCGGAA 597
 QY 21 SerGlyThrGluValLeuLeuIleAerSerArgProPheValGluTyraAnThrSer 40
 Db 598 AGTGAACGGAAAGAGTCTGCTAATGATGAGCGGCACTTTGGAAATACATACATCC 657
 QY 41 HsIleLeuGluAlaIleAenIleAenCysSerIleLeuMetIleValArgLeuGln 60
 Db 658 CACATTTTGGAGCCATTATATATCACTGTCACAGCTTGAAGGAGGTTGCAACAG 717
 QY 61 AspIleValLeuIleThrGluLeuIleGlnIleSerAlaIleHisIleValAspIleAer 80
 Db 718 GACAAAGTGTATATACAGAGCTCATCCAGCATTCAGCGAAATAGGTGACATTGAT 777
 QY 81 CysSerGluValValValValTyrAapGlnSerSerGlnAerValAlaSerLeuSerSer 100
 Db 778 TGCAGTCAGAGGTTGATGTTACGATCAAGCTCCCAAGATTTGCTCTCTTCA 837
 QY 101 AspCysPheLeuThrValLeuLeuGlyIleValLeuGluValSerPheAerValAlaLeu 120
 Db 838 GACTGTTTCTCACTGATCTTGGGTAACCTGGAAGAGCTTCAACTCTGTTCACTG 897
 QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuCysGlyIleValSer 140
 Db 898 CTTCGAGGTGGTGTGGTGTGCTCTGCTGTTGTTCCCTGGCCCTGTGTAAGGAAATCC 957
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAenIleGlyProThr 160
 Db 958 ACTTACGCTCCCTGACATTTCTGACCTTGTCTTACCTGTTCCACATTTGGGCAACC 1017
 QY 161 ArgIleLeuProAenLeuTyLeuGlyCysGlnArgValAlaAenIleValSerGlyLeu 180
 Db 1018 CGAATCTTCCCATCTTTATCTTGGCTGCGAGAGATGCTCTCAACAGAGCTGATG 1077
 QY 181 GlnGlnAenGlyIleGlyTyValLeuAenAlaSerTyThrCysProIleProAerPhe 200
 Db 1078 CAGCAGAAATGGATGGTGTATGTTAAATGACAGCATATCCCTGTCCAAAGCTGACTTT 1137
 QY 201 IleProGlnSerHisPheLeuArgValProValAlaAerSerPheCysGlyIleValLeu 220
 Db 1138 ATCCCGAGTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1197

QY 221 ProTleuAerPheSerValAerPheIleGlyValAlaValAerSerAenGlyCysVal 240
 Db 1198 CCGTGTGGACAAATTCAGTAAATTTCAATGAGAAAGCAAGCTTCAATGATGATGTT 1257
 QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyIleMet 260
 Db 1258 CTAGTCACTGTTTACCTGGATCTCCGCTCGGCACACATGGCTATGCCATACATG 1317
 QY 261 LysArgMetAerPheSerLeuAerGluAlaTyArgPheValIleGlyValArgProThr 280
 Db 1318 AAGAGATGACATGCTTTAGATGAGAGCTTACAGATTGTGAAAGAAAGAAAGACTACT 1377
 QY 281 IleSerProAerPheAerPheLeuGlyIleLeuAerTyGluValSerIleValAer 300
 Db 1378 ATATCTCCAACTTCATTTCTGGGCAACTCTGAGACTATGAGAAAGATTAAGAAC 1437
 QY 301 GlnThr 302
 Db 1438 CAGACT 1443

RESULT 3

AAAF30479
 ID AAF30479 standard; cDNA; 2118 BP.

AAAF30479;

29-MAY-2001 (first entry)

Human protein phosphatase and kinase protein-4 cDNA 1234795CB1.

Protein phosphatase and kinase protein; PPHK-4; human.

gastrointestinal disorder; immune system disorder; neurological disorder;
 cell proliferative disorder; cancer; diagnosis; therapy; ss.

Homo sapiens.

Key Location/Qualifiers

FT CDS 25..2025

FT misc_feature 163..222

FT misc_feature 1081..1140

FT /tag= b

FT /note= "unique fragment"

FT /tag= c

FT /note= "unique fragment"

FT /tag= c

FT /note= "unique fragment"

FT /tag= c

FT /note= "unique fragment"

FT /tag= c

FT /note= "unique fragment"

FT /tag= c

FT /note= "unique fragment"

FT /tag= c

FT /note= "unique fragment"

FT /tag= c

FT /note= "unique fragment"

FT /tag= c

CC tissues expressing PPKP-4) include gastrointestinal (0.385),
 CC cardiovascular (0.154), nervous (0.154) and reproductive (0.154).
 CC diseases or conditions associated with tissues expressing PPKP-4 (as a
 CC fraction of total tissues expressing PPKP-4) include cancer (0.692),
 CC inflammation or trauma (0.308) and cell proliferation (0.231). The
 CC encoded protein shows homology to mouse neuronal tyrosine threonine
 CC phosphatase 1. The invention provides human PPKP-1 to -11 polypeptides
 CC (see AAB20322-32) and polynucleotides (see AAF30476-86). It also provides
 CC expression vectors, host cells, antibodies, agonists and antagonists, as
 CC well as methods for diagnosing, treating or preventing disorders
 CC associated with expression of PPKP, including gastrointestinal
 CC disorders, immune system disorders, neurological disorders and cell
 CC proliferative disorders, including cancer

SO Sequence 2118 BP; 550 A; 552 C; 506 G; 510 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8,756-176 Length: 2118
 Score: 1552.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 0
 Query Match: 99.23% Indels: 0
 DB: 4 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AAF30479 (1-2118)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgPheValAlaLeuLeuGlu 20
 DB 25 ATGGCCCATGAGATGATGAGTCAATGTTACTGAGAGGTGGCTCGCTCGTGGAA 84
 QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluThrAsnThrSer 40
 DB 85 AGTGAACGGAAGAAAGCTGCTAATGATGAGCGGCAATTTGGAAATCAATCATCC 144
 QY 41 HisIleLeuGluAlaIleAsnIleAsnYSerIleuMetIleArgArgLeuGlnGln 60
 DB 145 CACATTTTGGAGACCATTAATATCAACTGCTCAAGCTTATGAAGCAAGTTGCAACG 204
 QY 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaValHisIleValAspIleAsp 80
 DB 205 GACAAAGTGTATTTACAGAGCTCATCCAGCATTCAGCAAAACATTAAGTGTGACATTGAT 264
 QY 81 CysSerGlnIleValValValValTyraAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB 265 TGGAGTCAGAGAGTGTGTAGTTCACATCAAGCTCCAAAGATGTTGCTCTCTCTTCA 324
 QY 101 AspCysPheLeuThrValLeuLeuGlyIleLeuGluIleYSerPheAsnSerValHisIleu 120
 DB 325 GACTGTTTCTACAGTACTTCTGGGTAACTGGAGAGAGACTTCAACTCTGTTCACTG 384
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyIleYSer 140
 DB 385 CTTCGAGGTGGGTGCTGAGTTCCTGCTGTTTCCCTGCTGTGAAAGAAATTC 444
 QY 141 ThrIleValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 445 ACTCTAGTCCCTACTGCTGATTTCTCAGCTGCTTACTCTGTTGCCAACATTTGGGCCAAC 504
 QY 161 ArgIleLeuProAsnLeuIleGlyCysGlnArgAspValLeuLeuIleGlyIleu 180
 DB 505 CGAATCTTCCCATCTTATCTTGCTGCTGAGGAGAGATGCTTCAACAGAGAGCTGAGG 564
 QY 181 GlnGlnAsnGlyIleGlyTyraValLeuAsnAlaSerTyraCysProIleProAspPhe 200
 DB 565 CAGAGAGATGGAGTGTATGTGTAAATGCCGCAATACCTGCTCAAAAGCTGACTTT 624
 QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
 DB 625 ATCCCGAGTCTCATTTCTGCGGTGCTGTAAAGACGCTTTTGTAGAAATTTTG 684
 QY 221 ProTyrLeuAspIleYSerValAspPheIleGluIleValAlaSerAsnGlyCysVal 240
 DB 685 CCGGTGTGGACAAATCAGTAGATTTCTAGAAAGCAAAAGCTTCCATGAGATGTGTT 744

QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
 DB 745 CTATGTCATCTTTTGTGCGGATCTCCGCTCCGCAACCAAGCTATCGCTATCATCATG 804
 QY 261 LysArgMetAspMetSerLeuAspGluAlaTyraArgPheValIleValGlyIleYAsnProThr 280
 DB 805 AAGAGAGATGACATGCTTTTATGATGAGACTTACAGATTTGTGAAGAAAGAAAGCTTACT 864
 QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluIleYAsn 300
 DB 865 ATATCTCCAACTTCAATTTTCTGCGCAACTCCTGACTATGAGAAAGATTAAGAAC 924
 QY 301 GlnThr 302
 DB 925 CAGACT 930
 RESULT 4
 AAD09492
 ID AAD09492 standard; DNA; 2732 BP.
 AC AAD09492;
 XX
 XX
 XX
 DT 10-SEP-2001 (first entry)
 DE Human SGP002 phosphatase polypeptide encoding DNA.
 KW Human; SGP002 phosphatase polypeptide; phosphatase-related disease;
 KW immune-related disorder; organ disease; organ transplant rejection;
 KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
 KW metabolic disorder; hematopoietic cancer; mood disorder; cardiac;
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
 KW attention disorder; cognition disorder; psychotic disorder; cytoskeletal;
 KW neurological disorder; virulence; nocotropic; cerebroprotective; therapy;
 KW neuroprotective; antibacterial; vulnery; tranquilliser; antistimatic;
 KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
 KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
 KW MRP; migraine; chromosome 12p11.1-p12.1; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 538..2535
 FT /tag=a
 FT /product="Human SGP002 phosphatase polypeptide"
 PN MO200146394-A2.
 XX
 PD 28-JUN-2001.
 XX
 PD 21-DEC-2000; 2000MO-US034736.
 PF
 XX
 PR 21-DEC-1999; 98US-0173255P.
 PR 28-DEC-1999; 99US-0175766P.
 PR 25-JAN-2000; 2000US-0178078P.
 PR 31-JAN-2000; 2000US-0179301P.
 XX
 PA (SUGEN-) SUGEN INC.
 XX
 PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;
 PI Hill RJ, Flanagan P;
 DR WPI; 2001-418058/44.
 DR P-PSDB; AAE04834.
 XX
 PT Novel phosphatase polypeptide useful for treating cancers, immune-related
 PT diseases and disorders, cardiovascular disease, brain or neuronal-
 XX associated diseases and metabolic disorders.
 XX
 PS Claim 29; Fig 1; 186pp; English.
 CC
 CC The present invention relates to phosphatase polypeptides, nucleotide

CC sequences encoding them, as well as various products and methods useful
 CC for the diagnosis and treatment of various phosphatase-related diseases
 CC and conditions. Substance that modulates the activity of phosphatase
 CC polypeptide is used to treat immune-related diseases and disorders,
 CC cardiovascular disease, brain or neuronal-associated diseases and
 CC metabolic disorder, including cancers of tissues, cancers of
 CC hematopoietic origin, diseases of central and peripheral nervous system,
 CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
 CC lateral sclerosis, viral infections, infections caused by prions,
 CC bacteria and fungi, ocular diseases, diabetes, migraine, pain, sexual
 CC dysfunction, mood disorders, attention disorders, cognition disorder,
 CC hypokinesia, hypertension, psychotic disorder, neurological disorders,
 CC dyskenasia and organ transplant rejection. The present sequence is a DNA
 CC encoding human SGP02 phosphatase polypeptide. This sequence is
 CC classified as dual specificity phosphatase (DSP) and MAP kinase
 CC phosphatase (MKP). SGP02 gene maps to chromosome 12p11.1-p12.1
 CC
 XX
 SQ Sequence 2732 BP; 710 A; 684 C; 686 G; 652 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1,3e-175 Length: 2732
 Score: 1552.00 Matches: 300
 Percent Similarity: 99.67% Conservative: 1
 Best Local Similarity: 99.34% Mismatches: 0
 Query Match: 99.23% Indels: 0
 DB: 4 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x AA009492 (1-2732)

QY 1 MetAlhIeGluNecIleGIYThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
 DB 538 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 597
 QY 21 SerGIYThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 DB 598 AGTGAACGGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 657
 QY 41 HisIleIleGluValIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 60
 DB 658 CACATTTTGGACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 717
 QY 61 AspIleValLeuIleThrGluLeuIleGlnIleSerAlaIleValIleValIle 80
 DB 718 GACAACTGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 777
 QY 81 CysSerGlnIleValIleValIleValIleValIleValIleValIleValIle 100
 DB 778 TGCAGTCAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 837
 QY 101 AspCysPheLeuThrValIleLeuGluValLeuGluValLeuGluValLeu 120
 DB 838 GACTGTTTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
 QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuGluGlyLeu 140
 DB 898 CTGGCAGGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAlaIleGlyProThr 160
 DB 958 ACTGAGTCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
 QY 161 ArgIleLeuProLeuLeuIleGluGluValLeuGluValLeuGluValLeu 180
 DB 1018 CGAATCTTCCCAATCTTAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
 QY 181 GlnGlnAlaGlnIleGlyIleValIleAlaAlaSerIleThrCysProIleProAspPhe 200
 DB 1078 CAGCAGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137
 QY 201 IleProGlnSerHisPheLeuArgValIleValAlaAlaSerPheCysGluValIleLeu 220
 DB 1138 ATCCCCGAGTCTCATTTCTGCGTGTCTGCTGATGATGATGATGATGATGAT 1197

QY 221 ProTrpLeuAspIleSerValIlePheIleGlyValIleAlaIleSerAlaGlyCysVal 240
 DB 1198 CCTGCTTGGACAAATCATTAATTTTATGAGAAAGCAAGCCCTCAATGATGATG 1257
 QY 241 LeuValIleCysLeuAlaGlyIleSerArgSerAlaIleAlaIleAlaIleIleMet 260
 DB 1258 CTAGTCACCTGTTAGCTGGATCTCCCTCCCTCCGACCATGCTATCGCTCAATCATG 1317
 QY 261 LysArgMetAspMetSerLeuAlaIleValIleValIleValIleValIleValIle 280
 DB 1318 AAGAGATGAGACATGCTTAACTGATGATGATGATGATGATGATGATGATGAT 1377
 QY 281 IleSerProAspPheAspPheLeuGluGlnLeuLeuAspTyrGluValIleValAsn 300
 DB 1378 AATTCCTCAATCTCAATTTTCTGGCCCAATCTCGATCATAGAAAGATTAAGAAC 1437
 QY 301 GlnThr 302
 DB 1438 CAGACT 1443

RESULT 5
 AAS15768
 ID AAS15768 standard; cDNA; 3059 BP.
 XX
 AC AAS15768;
 XX
 DT 18-TUN-2002 (first entry)
 XX
 DE cDNA sequence of human (dual specificity phosphatase) DUSP-10.
 XX
 KW Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke;
 KW neuronal degeneration syndrome; Alzheimer's disease; depression;
 KW schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
 KW osteoporosis; diabetes; gene therapy; chromosome 12; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 127..2124
 FT /crag= a
 FT /product= "DUSP-10 protein"
 XX
 PN MO200177340-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-EP003966.
 XX
 PR 10-APR-2000; 2000EP-00107143.
 XX
 PA (MERE) MERCK PATENT GMBH.
 XX
 PI Duecker K;
 XX
 DR WPI; 2002-010917/01.
 DR P-PSDB; AA009946.
 XX
 PT Novel dual specificity phosphatase polypeptides useful for treating
 PT cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
 PT disease, depression, schizophrenia, asthma and immune disorders.
 XX
 PS Claim 5; Page 34-37; 43pp; English.
 XX
 CC The present invention relates to a new isolated dual specificity
 CC phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
 CC sequence that is fully defined in the specification. The invention also
 CC provides a sequence encoded by a 3059 nucleotide sequence fully defined
 CC in the specification, and a sequence having at least 95 % identity to the
 CC polypeptide, or fragments or variants of DUSP-10. The invention is useful
 CC for treating cancer e.g. leukemia, colon carcinoma, lung cancer,
 CC prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
 CC stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
 CC schizophrenia, cardiac myotrophies, asthma, immune disorders,

CC inflammatory processes e.g. arthritis, bowel disease, type I diabetes,
CC osteoporosis, diabetes and diabetes associated diseases. The molecules of
CC the invention are also useful as vaccines for inducing immunological
CC response in a mammal, in disease diagnosis and in assays for screening
CC agonistic or antagonistic compounds. Other uses of the invention include
CC identifying membrane bound or soluble receptors, as a diagnostic reagent,
CC in chromosome localisation studies, and as a valuable tool in tissue
CC expression studies. The present sequence represents cDNA of the human
CC dual specificity phosphatase, DUSP-10. DUSP-10 is located on chromosome
CC 12

SQ Sequence 3059 BP; 831 A; 761 C; 709 G; 758 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.54e-175	Length:	3059
Score:	1552.00	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345a-109_COPY_1_302 (1-302) x AAS15768 (1-3059)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
DB 127 ATGGCCCATGAGATGATGAGTGAACCTCAATGTTTACGAGAGGTTGGCTGCTGCGGAA 186
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 187 AGTGAACGGAAGAAAGGCGCTAATTGATGAGCGGCGCTTTGTGATATCAATCATTC 246
QY 41 HisIleLeuGluAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 60
DB 247 CACATTTTGGAGCCCATTAATATCACTCTCCAGCTTATGAACGAAAGTTGCAACAG 306
QY 61 AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 307 GACAAAGTGTATTAACAGAGCTCATCCGATTCAGGCAACATTAAGGTTGACATTTGAT 366
QY 81 CysSerGluValValValValValValValValValValValValValValValVal 100
DB 367 TCGAGTCAGAAAGTGTATTAACATCAAGCTCCAAAGTGTGCTCTCTCTCTCTCA 426
QY 101 AspCysPheLeuThrValLeuLeuGlyLeuLeuGlyLeuLeuLeuLeuLeuLeuLeu 120
DB 427 GACTGTTTCTCACTGTAATCTTGGGTAACCTGGAGAAAGCTTCAACTGTTCACTG 486
QY 121 LeuAlaGlyValPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlySer 140
DB 487 CTTCGAGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 546
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAlaIleGlyProThr 160
DB 547 ACTCTAGTCCCTCACTGCAATTTCTCAAGCTTGTACTCTGTGCAACATTTGGGCAAC 606
QY 161 ArgIleLeuProAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
DB 607 CGAATTTCTCCCAATCTTATCTTGGGCGCCAGCAAGATGCTTCAACAAAGAGCTGAT 666
QY 181 GlnGlnGlnGlyIleGlyValValLeuAlaSerThrCysProGlyProAspPhe 200
DB 667 CAGCAGAAATGGAGTGTATGATGTTAAATGCCAGCAATCTGTCCAAAGCTGACTTT 726
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlyValIleLeu 220
DB 727 ATCCCCGAGTCTCATTTCTGCGTGTGCTGTGATATGACGCTTTGTGGAAGAAATTTG 786
QY 221 ProTrpLeuAspIleValValValValValValValValValValValValValVal 240
DB 787 CCGGTGTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCTCCATGATGTGTT 846
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAla 260

DB 847 CTATGCACTGTTAGTGGATCTCCGCTCCGCCAACATCGCTATGCTATCATCATG 906
QY 261 LysArgMetAspMetSerLeuAspGluAlaValTrpArgPheValValValValValVal 280
DB 907 AAGAGATGAGCAGATGCTTTTGAATGAGCTTACGATTTTGAAGAAAAAGACCTAAT 966
QY 281 IleSerProAsnPheAsnPheLeuGlyGluLeuLeuAspTrpGlyValValValValVal 300
DB 967 ATATCTCCAACTCAATTTTCTGGCCAACTCTGAGCTATGAGAAAGATTAAAGAAC 1026
QY 301 GlnThr 302
DB 1027 CAGACT 1032

RESULT 6
ABN59704
ID ABN59704 standard; cDNA; 3104 BP.
XX
XX
XX
AC ABN59704;
XX
XX
DT 28-JUN-2002 (first entry)

XX
XX
DE Novel human coding sequence SEQ ID NO: 115.

XX
KW Human; anti-naemic; vulnerary; anti-inflammatory; immunomodulator;
KW anti-infectivity; cerebroprotective; cytosolic; rheumatic; gene therapy;
KW neuroprotective; anti-Parkinsonian; protein therapy; ESR;
KW expressed sequence tag; gene; ss.

OS Homo sapiens.

PN WO200222660-A2.

XX
XX
PD 21-MAR-2002.

XX
XX
PF 10-SEP-2001; 2001MO-US026015.

XX
XX
PR 11-SEP-2000; 2000US-00659671.

XX
XX
PA (HYSE-) HYSEQ INC.

XX
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX
XX
DR WPI; 2002-292408/33.

XX
XX
DR P-PDB; ABB97291.

PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.

XX
XX
PS Claim 1; SEQ ID NO 115; 509bp; English.

XX
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to help tissue regrowth
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention

SQ Sequence 3104 BP; 845 A; 772 C; 718 G; 769 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.58e-175	Length:	3104
Score:	1552.00 <td>Matches:</td> <td>300</td>	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345A-109_COPY_1_302 (1-302) x ABN59704 (1-3104)

```

QY      1 MetAlAHISgluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB      159 ATGGCCCATGATGATGGAACTCAAAATGTTACTGAGAGGTTGGCTGCTCGTGGAA 218
QY      21 SerGlyThrGluArgValLeuLeuIleAspSerArgProPheValGluValPheThrSer 40
DB      219 AGTGGAAAGGAAAGAGTCTGCTGCTATGATGACCGGCGCATTTGGGAAATCAATACATCC 278
QY      41 HisIleuGluIleAlaIleAsnIleAsnCySerIleuMetIleuArgArgLeuGln 60
DB      279 CACATTTGGAGACGATTAATATCACTGCTCCAGAGTTTGAAGAGAGTTGCAACAG 338
QY      61 AspValValLeuIleThrGluLeuIleGlnHisSerAlaValIleValIleAsp 80
DB      339 GACAAAGTTTAAATTACAGAGCTATCCAGCATTCACGAAACATTAAGTTGACATTGAT 398
QY      81 CySerGlnIleValValIleValIleAspGlnSerSerGlnAspValIleSerIleuSer 100
DB      399 TGCAGTCAGAGGTTGATGTTACGATCAAGCTCCAAAGATGTTGCTCTCTCTTCA 458
QY      101 AspCyPheLeuThrValIleuLeuGlyLeuGluIleuSerPheAsnSerValIleLeu 120
DB      459 GACGTTTCTCACTGATCTTGGGTAACCTGGAGAGAGCTTCACTGTTCACTG 518
QY      121 LeuAlaGlyIlePheAlaGluPheSerArgCyPheProGlyLeuCyGluGlyIleuSer 140
DB      519 CTTCAGAGTGGGTTGCTGATCTCTGTTGTTCCCTGCTGCTGTAAGAGAAATCC 578
QY      141 ThrLeuValProThrCysIleSerGlnProCysLeuProValIleAlaIleGlyProThr 160
DB      579 ACTCTAGTCCCTACCTGATTTCTCAAGCTTGTCTTACTGTCGCAACATTTGGCCAAAC 638
QY      161 ArgIleuProAsnLeuIleuGlyCysGlnArgAspValIleuLeuIleuGlyIleuLeu 180
DB      639 CGAATCTTCCCAATCTTATCTTGGCTGCTGACAGAGATGCTCTCAACAGAGAGCTGATG 698
QY      181 GlnGlnAsnGlyIleGlyIleValIleuAsnIleSerIleuValIleuValIleuValIleu 200
DB      699 CACGAGATGGATGGTATGTTATGTTAAATGCCAGCAATCTGTCAAAGCTGACTTT 758
QY      201 IleProGlnSerIlePheLeuArgValProValAsnAspSerPheCysGlyIleuVal 220
DB      759 ATCCCGAGTCTCAATCTTCTGCGTGGCTGCTGATGACAGCTTTGTGAGAAATTTTG 818
QY      221 ProTrpLeuAspIleSerValIlePheIleGlyIleValIleValIleValIleValIle 240
DB      819 CCCTGCTTGAACAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 878
QY      241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
DB      879 CTATGTCACCTGTTAGCTGGATCTCCGCTCGCCACATCTGCTATCGCTACATCATG 938
QY      261 LysArgMetAspMetSerIleuAspGluAlaIleValIleValIleValIleValIle 280
DB      939 AAGAGATGAGCATGCTTTAGATGAGAGCTTACAGATTGTGAAAGAAAAAACCCTACT 998
QY      281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIleValIleValIleValIle 300
DB      999 ATATCTCCCAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAGAGATTAAGAAC 1058
QY      301 GlnThr 302
DB      1059 CAGACT 1064

```

RESULT 7

ABK47596
ID ABK47596 standard, cDNA, 3496 BP.
AC ABK47596;
XX

DT 02-JUL-2002 (first entry)

DE XX cDNA encoding human dual-specificity phosphatase-3 (DSP-16) protein.

XX Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
XX mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
XX cancer; graft-versus-host disease; allergy; metabolic disease;
XX abnormal cell growth; abnormal cell proliferation; contact inhibition;
XX cell cycle abnormality; anchorage independent cell growth; apoptosis;
XX intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; 89.

XX Homo sapiens.

OS XX

FT Key Location/Qualifiers

FT CDS 562..2559

FT /tag= a

FT /product= "Human dual-specificity phosphatase-3 (DSP-16)

PN WO200226997-A2.

XX 04-APR-2002.

XX 25-SEP-2001; 2001WO-US030124.

XX 26-SEP-2000; 2000US-0235487P.

XX (CEPT-) CEPTYR INC.

XX Luche RM, Wei B;

XX WPI: 2002-315802/35.

XX P-PSDB; AAU79156.

XX New DSP-16 polypeptide, useful for identifying modulators of its

XX PT activity, which can be used in the treatment of disorders such as

XX Duchenne muscular dystrophy, or cancer.

XX Claim 7, Fig 1; 87p; English.

XX The present invention relates to a new polypeptide, DSP-16, having a 665

XX amino acid sequence, given in the specification, or a variant having at

XX least 50 % identical residues, which retains the ability to

XX dephosphorylate an activated mitogen-activated protein (MAP) kinase. The

XX invention can be used for identifying agents which modulate DSP-16

XX activity, for modulation of a proliferative response in a cell, survival

XX of a cell, or differentiation of a cell. The cell displays contact

XX inhibition of cell growth or anchorage independent growth and may display

XX altered intercellular adhesion. The agent may modulate apoptosis, or the

XX cell cycle. The identified modulators can be used to treat Duchenne

XX muscular dystrophy, cancer, graft-versus-host disease, autoimmune

XX diseases, allergies, metabolic diseases, abnormal cell growth, abnormal

XX cell proliferation, and cell cycle abnormalities. The present nucleic

XX acid sequence is that of the human dual-specificity phosphatase-3 (DSP-

XX 16) gene located on chromosome 12p. This sequence encodes the human DSP-

XX 16 protein of the invention

SQ Sequence 3496 BP; 939 A; 854 C; 856 G; 847 T; 0 U; 0 Other;

Alignment Scores:

Score: 1 9e-175 Length: 3496

Percent Similarity: 1552.00 Matches: 300

Best Local Similarity: 99.67% Conservative: 1

Query Match: 99.34% Mismatches: 1

DB: 99.23% Indels: 0

Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x ABK47596 (1-3496)

QY 1 MetAlAHISgluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 562 ATGGCCCATGATGATGGAACTCAAAATGTTACTGAGAGGTTGGCTGCTCGTGGAA 621

QY	21	SerGlyThrGluValLeuLeuLLeaSpSerAProPheValGluTyrAsnThrSer	40
Db	622	AGTGAAGAGAAAAAGCTGCTTAATGATGACCGGCGCATTTGGAAATCAATACATCC	681
QY	41	HisIleLeuGluAlaIleAsnIleAsnCYSerSerIysLeuMetIysArgAglLeuGln	60
Db	682	CACATTTTGGAAAGCATTAATATCACTGCTCCAAAGCTTATGAAGCGAAGGTGCACAG	741
QY	61	AspIysValLeuIleThrGluLeuIleGlnHisSerAlaIysHisIysValAspIleAsp	80
Db	742	GACAAAGGTAAATTAACAGAGCTCATCAGCATTCACAGCAACATAGTTGACATTGAT	801
QY	81	CysSerGlnIysValIleValItyrAspGlnSerSerGlnAspValIleSerLeuSerSer	100
Db	802	TGCAGTCAGAAAGGTGTATGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA	861
QY	101	AspCysPheLeuThrValIleLeuGluIlyAsLeuGluIlySerPheAsnSerValHisLeu	120
Db	862	GACGTGTTTCTCACTGTACTTCTGGGTAACTGGAGAAAGCTTCAACTCTGTACCTGG	921
QY	121	LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyIleuCYseGluGlyIysSer	140
Db	922	CTTGCAAGTGGGTGTTGCTGAGCTTCTCTGTTGTTCCCTGCGCTGTGAGAAATATCC	981
QY	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	982	ACTCTAGTCCCTCACTGACATTTCTCAGAGCTTGCTTACTGTTGCCAAATTTGGCCAAAC	1041
QY	161	ArgIleLeuProAsnLeuTyrIleuGlyCysGlnArgAspValLeuAsnIyGluLeuIle	180
Db	1042	CGAATCTTCCCAATCTTATCTTGCTGCGCAGCGAGATGTCCTCAACAGAGAGCTGATG	1101
QY	181	GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProIysProAspPhe	200
Db	1102	CAGCAGAATGGAGTGTGTTATGTGTTAAATGCCAGAAARCTGTCCAAAGCTGACCTT	1161
QY	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIysIleLeu	220
Db	1162	ATCCCGAGATCTCAATTTCTGCGGTGCTGTGAATGACAGCTTTGTGACAAAATTTTG	1221
QY	221	ProTyrLeuAspIysSerValAspPheIleGluIySAIAlaIySAIAserAsnIyCysVal	240
Db	1222	CCGTGGTTGCAAAATCAATACATTAATTTCAATGGAAAGCAAAAGCTCCAAATGAGATGTT	1281
QY	241	LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet	260
Db	1282	CTAGTGCACTGTTTGTGCTGGGATCTCCCGCTCCGCCACATGCGTATGCGCTACATCATG	1341
QY	261	IysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIyGluIyAsnArgProThr	280
Db	1342	AAGAGAGATGAGATGTCTTATGATAGACTTTCACATTTGTGAAAGAAAAAGACTTACT	1401
QY	281	IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluIyIySAIleIySAIAsn	300
Db	1402	ATATCTCAAACTTCATATTTCTGGGGCAACTCTCGACTATGAGAGAAATTTAAGAC	1461
QY	301	GlnThr 302	
Db	1462	CAGACT 1467	
RESULT 8			
AA514639			
ID	AA514639 standard, cDNA, 3544 BP.		
XX	AA514639;		
XX			
XX	18-DEC-2001 (first entry)		
XX			
DE	Human cDNA encoding dual specificity phosphatase 21117.		
FW	Human, sex, dual specificity phosphatase 21117, hepatotropic; cytosolic;		
FW	hematopoietic disorder; autoimmune disorder; diabetes mellitus;		
FW	rheumatoid arthritis; multiple sclerosis; Crohn's disease;		

KM	liver disorder; erythroid associated disorder; hemolytic anaemia;
KM	cellular proliferative; differential disorder; leukaemia;
KM	acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer.
XX	
OS	Homo sapiens.
XX	
XX	Key
FT	CDS
FT	Location/Qualifiers
FT	589. .2586
FT	/tag=a
FT	/product= "Dual specificity phosphatase 21117"
XX	
XX	WO200173059-A2.
XX	
XX	04-OCT-2001.
XX	
XX	23-MAR-2001; 2001WO-US009477.
XX	
XX	24-MAR-2000; 2000US-0191858P.
XX	
XX	(MILL-) MILLENNIUM PHARM INC.
XX	
XX	Meyers RA;
XX	
XX	WPI; 2001-611635/70.
XX	
XX	P-PSDB; AAU09016.
XX	
XX	New human dual specificity polypeptides and nucleic acids for diagnosis
PT	of disease and treatment of e.g. liver disorders.
PT	
PT	Claim 1; Fig 1; 143pp; English.
XX	
XX	The invention relates to two novel human dual specificity phosphatases
CC	designated 21117 and 38692, the nucleic acids encoding them (including
CC	fragments, allelic variants, their complements or nucleic acids that
CC	hybridise to them) and antibodies raised against the proteins. The
CC	antibody is useful for detecting the presence of the polypeptide, and the
CC	nucleic acid fragments are useful for detecting the presence of the
CC	nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
CC	antisense sequences) are useful for modulating the activity or expression
CC	of the polypeptide/nucleic acid, useful for the treatment of e.g.
CC	hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
CC	rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
CC	listed in the specification), liver disorders, erythroid associated
CC	disorders (e.g. haemolytic anaemia) cellular proliferative or
CC	differential disorders, leukaemias (e.g. acute myeloid leukaemia),
CC	carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and
CC	38692 are also useful for modulating the proliferation, survival, the
CC	migration or differentiation of a 38692 or 21117-expressing cell. The
CC	polypeptide and nucleic acids are useful for identifying modulating
CC	agents. The present sequence encodes the dual specificity phosphatase
XX	21117
XX	
XX	Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;
XX	
XX	Alignment Scores:
XX	
XX	Pred. No.: 1 94e-175 Length: 3544
XX	Score: 1552.00 Matches: 300
XX	Percent Similarity: 99.67% Conservative: 1
XX	Best Local Similarity: 99.34% Mismatch: 1
XX	Query Match: 99.23% Indels: 0
XX	DB: 5 Gaps: 0
XX	
XX	US-10-029-345A-109_COPY_1_302 (1-302) x AAS14639 (1-3544)
XX	
QY	1 MetAlHISgluMetIleGIYThrglnIleValThrgluArgLeuValAlaLeuEngIn 20
DB	589 ATGGCCCAATGATGATATGGAACTCAATTGTCTAGAGAGTGGGGCTCGCTGGAA 648
QY	21 SerGIYThrgluValLeuLeuIleAspSerArgProPheValGluIlyrAsnThrSer 40
DB	649 AGTGGACGGAAGAAAGTCTGCTAAATGATAGCCGGCCATTGTGGAAATACATATCC 708
QY	41 HsIleLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 60

[illegible]

FT	1	US2002034807-A1.	/*tag= a	/product= "human dual specificity phosphatase 21117"	
FT	2			/note= "Specifically claimed in claim 2"	
FT	3				
XX	4				
XX	5				
XX	6				
XX	7				
XX	8				
XX	9				
XX	10				
XX	11				
XX	12				
XX	13				
XX	14				
XX	15				
XX	16				
XX	17				
XX	18				
XX	19				
XX	20				
XX	21				
XX	22				
XX	23				
XX	24				
XX	25				
XX	26				
XX	27				
XX	28				
XX	29				
XX	30				
XX	31				
XX	32				
XX	33				
XX	34				
XX	35				
XX	36				
XX	37				
XX	38				
XX	39				
XX	40				
XX	41				
XX	42				
XX	43				
XX	44				
XX	45				
XX	46				
XX	47				
XX	48				
XX	49				
XX	50				
XX	51				
XX	52				
XX	53				
XX	54				
XX	55				
XX	56				
XX	57				
XX	58				
XX	59				
XX	60				
XX	61				
XX	62				
XX	63				
XX	64				
XX	65				
XX	66				
XX	67				
XX	68				
XX	69				
XX	70				
XX	71				
XX	72				
XX	73				
XX	74				
XX	75				
XX	76				
XX	77				
XX	78				
XX	79				
XX	80				
XX	81				
XX	82				
XX	83				
XX	84				
XX	85				
XX	86				
XX	87				
XX	88				
XX	89				
XX	90				
XX	91				
XX	92				
XX	93				
XX	94				
XX	95				
XX	96				
XX	97				
XX	98				
XX	99				
XX	100				
XX	101				
XX	102				
XX	103				
XX	104				
XX	105				
XX	106				
XX	107				
XX	108				
XX	109				
XX	110				
XX	111				
XX	112				
XX	113				
XX	114				
XX	115				
XX	1				

```
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGluLeuCysGluGlyLysSer 140
DB 949 CTGGACAGTGGGTTGCTGAGATCTCTCGTGTGTTTCCCTGGCTCTGTGAAGGAAATCC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 1009 ACTCTAGTCCTTACTGATTTCTCAGCCTTGCTTACTGTTGGCAACATTGGGCCAAC 1068
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgPheValLeuAsnIleGluLeu 180
DB 1069 CGAATTCCTCCCAATCTTATCTTGCTGGCCAGAGATGTCCTCAACAAAGACTATG 1128
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
DB 1129 CAGAGAGATGGGATGTTAGTGTATATCCAGCAATACCTGTCCAAAGCTGACTTT 1188
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
DB 1189 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGATGACAGCTTTTGTGAAGAAATTTTG 1248
QY 221 ProTyrLeuAspLysSerValAspPheIleGluLysAlaLysSerAsnGlyCysVal 240
DB 1249 CCGTGTGTGGACAATCAGTATGATTTCATTGAGAAAGCAAGCCCTCCAAATGGATGTGTT 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
DB 1309 CTAGGCACTGTTTACTGGGATCTCCCGCTCCGCCACCACTGCTATGCTTACTATCATG 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
DB 1369 AAGAGGATGACATGCTTTAGATGACAGCTTACAGATTGTGAAAGAAAGAAAGACTTACT 1428
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysValIleLysAsn 300
DB 1429 ATATCTCCAAACTTCAAATTTCTGGGCCAACTCTCGACATATGAGAGAGATTAAAGAC 1488
QY 301 GlnThr 302
DB 1489 CAGACT 1494
RESULT 10
ABK14474 ID ABK14474 standard; cDNA; 3766 BP.
AC ABK14474;
DT 08-MAY-2002 (first entry)
DE Human protein phosphatase 7 (PP7) cDNA sequence.
XX Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
XX acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
XX Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
XX dementia; Parkinson's disease; developmental disorder; Down's syndrome;
XX cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
XX melanoma; myeloma sarcoma; gene; ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 538..2535
FT /tag= a
FT /product= "Protein_phosphatase_7 (PP7) "
XX
XX WO200210363-A2.
XX
XX 07-FEB-2002.
XX
XX 26-JUL-2001; 2001WO-US023716.
XX
XX 28-JUL-2000; 2000US-0221679P.
XX
XX 03-AUG-2000; 2000US-0223272P.
XX
XX 10-AUG-2000; 2000US-0224309P.
```

```
PR 18-AUG-2000; 2000US-0226728P.
PR 30-AUG-2000; 2000US-0229254P.
PR 08-SEP-2000; 2000US-0231366P.
XX
XX (INCY- ) INCYTE GENOMICS INC.
XX
XX Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE,
PI Stewart BA, Gandhi AR, Patterson C, Lee EA, Hatfield AJA, Lu DM,
PI Tribouley CM, Griffen JB, Baughn MR, Yue H, Warren BA, Nguyen DB,
PI Malia NK, Kearney L;
XX
XX WPI; 2002-188735/24.
DR P-PSDB; AAU75789.
XX
XX New protein phosphatases, useful for diagnosing, treating or preventing
PT immune system disorders (e.g. Crohn's disease), neurological disorders
PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
PT cancers).
XX
XX Claim 5; Page 114-115, 117pp; English.
```

The present invention relates to a new polypeptide, a naturally occurring amino acid sequence at least 95 % identical to it, a biologically active fragment of it or an immunogenic fragment of it. The polypeptides, polynucleotides, agonists and antagonists are useful for diagnosing, treating or preventing disorders associated with aberrant expression of protein phosphatases (PP), particularly immune system disorders e.g. acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia, asthma or Crohn's disease, neurological disorders e.g. epilepsy, Huntington's disease, dementia or Parkinson's disease, developmental disorders e.g. Down's syndrome, or cell proliferative disorders e.g. cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma or sarcoma. The present nucleic acid sequence encodes human protein phosphatase 7 (PP7) which is one of several human protein phosphatases (AAU75783-AAU75792) of the invention

Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:
2.13e-175	3766	1552/00
Score: 1552.00		
Percent Similarity: 99.67%		Conservative: 1
Best Local Similarity: 99.34%		Mismatches: 1
Query Match: 99.23%		Indels: 0
DB: 6		Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x ABK14474 (1-3766)

```
QY 1 MetalahisgluWectilegIYThrglnlleValThrcIuArgLeuValAlaLeuGlu 20
DB 538 ATGGCCCATGAGATGGATGGAACTCAAAATGTTACTGAGAGGTTGGGCTGTGGAA 597
QY 21 SerGlyThrglnLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
DB 598 AGTGAACGGAAGAAAGTCTGCTTAATGATAGCCGGCACTTGTGGAAATCAATATCC 657
QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerLysLeuWectIysArgIleGlnGln 60
DB 658 CACATTTTGGAAAGCCATTAATATATCACTGCTCCAAAGCTTATGAGCAAGGTTGCAACAG 717
QY 61 AspLysValLeuIleThrglnLysIleGlnHisSerLysValLysValAspIleAsp 80
DB 718 GACAAAGTGTAAATTACAGAGCTATCCAGCATTCAGAGAACTTAAGTTGCTCTTCA 777
QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 778 TGCAGTCAGAAAGTTGTAGTTTACGATCAAAAGCTCCCAAGATGTGCTCTCTTCA 837
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeu 120
DB 838 GACTGTTTCTCACCTGATCTCTGGTAAACTGAGAGAAAGCTTCAACTCTGTTCACCTG 897
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGluLeuCysGluGlyLysSer 140
```



```
DB 898 CTTGCGAGTGGGTTTCTGAGTCTCTGTTGTTTCCCTGCTGTGAGGAAATCC 957
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaIleAsnIleGlyProThr 160
DB 958 ACTGTGTCCTTACCTGACATTTCTCAGCCCTGCTTACCTGTTCCAAATGGGCCCAACC 1017
QY 161 ArgIleLeuProAsnLeuTyrlleuGlyCysGlnArgAspValLeuAsnIleGlyLeu 180
DB 1018 CGAATCTTCCCAATCTTATCTTGCTGCGACGAGAGTCTCTCAACAGAGGCTGATG 1077
QY 181 GlnGlnAsnGlyIleGlyTyrlleuValLeuAsnAlaSerTyrlleuCysProIleProAspPhe 200
DB 1078 CAGCAAGATGGGATTGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 1137
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluTyrlleu 220
DB 1138 ATCCCGAGCTCATTTCTGCGCTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 1197
QY 221 ProTrpLeuAspLysSerValAspPheIleGluValAlaIleAsnIleGlyCysVal 240
DB 1198 CCGTGGTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCTCCAAATGGATGTGT 1257
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleuMet 260
DB 1258 CTAGTCACTGTTTACCTGGGATCTCCGCTCCGCAACCATCCCTATCCCTACATCATG 1317
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrlleuArgPheValIleGluLysArgProThr 280
DB 1318 AAGAGAGATGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGACCTACT 1377
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrlleuValIleLysAsn 300
DB 1378 AATCTTCCAAACTTCAATTTCTGGGCCAACTCTGAGACTATAGAGAAAGATTAAAGAAC 1437
QY 301 GlnThr 302
DB 1438 CAGACT 1443

RESULT 11
ABN83966
ID ABN83966 standard; DNA; 4790 BP.
XX
AC ABN83966;
XX
DT 06-SEP-2002 (first entry)
XX
DE Human gene sequence #13.
XX
KM Human; brain; tonsil; hippocampus; foetal brain; diagnosis; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 184..2181
FT /tag= a
XX
PN WO200252005-A1.
XX
PD 04-JUL-2002.
XX
PF 20-DEC-2001, 2001MO-JP011217.
XX
PR 22-DEC-2000, 2000JP-00389742.
XX
PA (KAZU-) KAZUSA DNA RES INST FOUND.
PA (CELE-) CELESTAR LEXICO-SCI LTD.
XX
PI Ohara O, Nagase T, Nakajima D;
XX
DR MPI; 2002-500762/53.
XX
P-PSDB; ABB97946.
```

```
PT Genes and their expression products cloned from human cDNA libraries for
FT treatment and diagnosis of diseases associated with their expression.
XX
XX Claim 1(a); Page 111-117; 238pp; Japanese.
XX
CC The invention relates to DNA encoding polypeptides directly cloned from
CC cDNA libraries originating in adult whole brain, human tonsil, human
CC adult hippocampus and human foetal whole brain. Polypeptides and
CC polynucleotides of the invention may be used in the investigation of
CC differential expression of the DNA sequences in normal subjects and
CC disease patients. They may also be used in the production of antibodies,
CC oligonucleotide probes and DNA chips for diagnosis and identification of
CC drugs for treatment of diseases with which the DNA sequences are
CC associated. The sequences given in records ABN83954-ABN83984 represent
CC human gene sequences of the invention
XX
SQ Sequence 4790 BP; 1293 A; 1064 C; 1152 G; 1281 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3, 08e-175 Length: 4790
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 6 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x ABN83966 (1-4790)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
DB 184 ATGGCCCATGAGATGATGGAATCACTCAAAATTTGTTACTGAGAGGTGTGCTGTGAA 243
QY 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrlleuThrSer 40
DB 244 AGTGAACCGAAAGAGTGTGCTTAATGTATGATACCGGCTATTTGTGAATCAATACATCC 303
QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerLysLeuMetLysArgArgLeuGln 60
DB 304 CACATTTTGGAGGCCATTAAATCACTCTCCAACTTATGAGGAGGTTGCAACAG 363
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
DB 364 GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCAACATAGAGTTGACATTGAT 423
QY 81 CysSerGlnLysValValIleValIleAspGlnSerSerGlnAspValAlaSerLeuSer 100
DB 424 TGCAGTCAGAGAGTTGTAGTTACGATCAAGCTCCCAAGATGTTCCCTCTCTCTTCA 483
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
DB 484 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGCTTCAACTGTTCACCTG 543
QY 121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuGlyGluGlyLysSer 140
DB 544 CTTGCGAGTGGGTTTGTGAGTCTCTCGTTGTTTCCGGGCTCTGTGAGGAAATCC 603
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaIleAsnIleGlyProThr 160
DB 604 ACTGTGTCCTTACCTGACATTTCTCAGCCCTGCTTACCTGTTGCCAAATGGGCCCAACC 663
QY 161 ArgIleLeuProAsnLeuTyrlleuGlyCysGlnArgAspValLeuAsnIleGlyLeu 180
DB 664 CGAATCTTCCCAATCTTATCTTGCTGCGACGAGAGTCTCTCAACAGAGGCTGATG 723
QY 181 GlnGlnAsnGlyIleGlyTyrlleuValLeuAsnAlaSerTyrlleuCysProIleProAspPhe 200
DB 724 CAGCAAGATGGGATTGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 783
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluTyrlleu 220
DB 784 ATCCCGAGCTCATTTCTGCGCTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 843
QY 221 ProTrpLeuAspLysSerValAspPheIleGluValAlaIleAsnIleGlyCysVal 240
```


DB 844 CCGTGGTGCACAAATCAGTATTTTCATTGAGAAAGCAAAAGCCCTCCCATGATGTCTT 903
QY 241 LeuValHisCysLeuAlaIleGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
DB 904 CTAGTGCACTGTTTACCTGGATCTCCCGCTCCGACCATGCTATGCGCTACATCATG 963
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluLysArgProThr 280
DB 964 AAGAGGATGACATGTCTTATGATGAGCTTACAGATTTTGTAAGAAAAGAAAGACCTTCT 1023
QY 281 ILeSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
DB 1024 ATATCTCCAAACTCAATTTTCTGGGCCCAACTCTGACCTATGAGAAAGATTAAAGAAC 1083
QY 301 GlnThr 302
DB 1084 CAGACT 1089
RESULT 12
ABV20833
ID ABV20833 standard; cDNA; 5145 BP.
XX
AC ABV20833;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 20824.
XX
KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JB;
PI
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3419; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 3,44e-175 Length: 5145
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 5 Gaps: 0
US-10-029-345a-109_copy_1_302 (1-302) x ABV20833 (1-5145)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGlnArgLeuValAlaLeuGln 20
DB 589 ATGGCCCATAGATGATGTAAGTCAATTTGTTACTGAGAGGTGGCTCTGCTGAA 648
QY 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
DB 649 AGTGAACGGAAAAAGTGGCTGTAATTGATAGCCGGCCATTGTGGAATACATCATCC 708
QY 41 HisIleLeuGlnAlaIleAsnIleAsnGlySerIleLeuMetIleValArgArgLeuGln 60
DB 709 CACATTTGGAAAGCCATTAATATCACTGCTCAAGCTTATGAAAGGATTGCAACG 768
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaIleHisLysValAspIleAsp 80
DB 769 GACAAAGTGTATTTACAGAGCTCATCCAGCATTCAGCAAAACATAGGTTGACATTGAT 828
QY 81 CysSerGlnLysValValValIleValIleAspGlnSerSerGlnAspValAlaSerLeuSer 100
DB 829 TGCAGTCCAGAGGTTGATGTTATACGATCAAACTCCCAAGATGTTCCTCTCTTCA 888
QY 101 AspCysPheLeuThrValLeuLeuGlyLysGluLysSerPheAsnSerValHisLeu 120
DB 889 GACTGTTTCTCACTGACTTCTGGGTAACTGGAGAAAGCTTCAACTGTTCACCTG 948
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
DB 949 CTTCGAGGTGGGTTTGCTGAGTTCTCTGTTTCCCTGGCTCTGTAAGAAATTC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 1009 ACTCTAGTCCATCCATCGATTCCTCAGCCCTCTTACCTGTTGCCAAATGGGGCAACC 1068
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
DB 1069 CGAATTTCTCCCAATCTTTATCTTGCTGCCAGCAAGATGCTCTCAAGAGGCTGATG 1128
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
DB 1129 CAGCAGAAATGGGATGTTATGTATGTTAAATGCCAGCAATACCTGCCAAAGCTGACTTT 1188
QY 201 ILeProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleLeu 220
DB 1189 ATCCCGAGATTCATTTTCCGTGGTGCTTGTAATGACACTTTTGGAAAAATTTTG 1248
QY 221 ProTyrPheAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
DB 1249 CCGTGGTTGGACAAATCATGATTTCTTGAAGAAAGAAAGCTCCAAATGATGTCTT 1308
QY 241 LeuValHisCysLeuAlaIleGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
DB 1309 CTAGTGCACTGTTTACCTGGATCTCCCGCTCCGACCATGCTATGCTCATCATG 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluLysArgProThr 280
DB 1369 AAGAGGATGACATGTCTTATGATGAGCTTACAGATTTTGAAAGAAAAGAAAGCTCACT 1428
QY 281 ILeSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAsn 300
DB 1429 ATATCTCCAAACTCAATTTTCTGGGCCCAACTCTGACCTATGAGAAAGATTAAAGAAC 1488
QY 301 GlnThr 302
|||||

DB 1489 CAGACT 1494

RESULT 13

ABV21080

ID ABV21080 standard; cDNA; 5145 BP.

XX

AC ABV21080;

XX

DT 13-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 21071.

XX

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX OS

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189662P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX

PI Schlegel R, Endege WO, Monahan JE;

XX

DR WPI; 2001-662795/76.

XX

PT Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful

PT for detecting presence of prostate cancer, stage of prostate cancer.

XX

XX

PS Claim 1; Page 3481; 11750bp; English.

XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for: (a) assessing whether

CC a patient is afflicted with prostate cancer; (b) monitoring the

CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)

CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX

SO Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Alignment Scores:

Pred. No. 1 3.44e-175 Length: 5145

Score: 1552.00 Matches: 300

Percent Similarity: 99.67% Conservative: 1

Best Local Similarity: 99.34% Mismatches: 1

Query Match: 99.23% Indels: 0

DB: 5 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x ABV21080 (1-5145)

QY 1 MetAlHieGluMetIleGlyThrGlnIleValThrGluArgIleValAlaLeuGln 20

DB 589 ATGGCCATGAGATGATGGAACCTCAATTGTTACTGAGGCTTGCGCTCTGCGAA 648

QY 21 SerGlyThrGlnIleValLeuLeuIleAspSerArgProPheValGlnTyrAenThrSer 40

DB 649 AGTGAACGAAAAAGTGCTGCTAATTGATAGCCGCGCATTTGTGSAATACATATACCC 708

QY 41 HieIleLeuGlnAlaIleAsnIleAsnCySerIleLeuMetIleValArgIleGlnGln 60

DB 709 CACATTTTGGAAAGCCATTATATATCACTGCTCCMACCTTATAGAGGAAGGTTGCAACG 768

QY 61 AspIleValIleuIleThrGluLeuIleGlnIleSerAlaIleHisIleValAspIleAsp 80

DB 769 GACAAAGTTTATTTATTCAGAGCTCATCCAGCTTCCAGCAAACTAAGTTGACATTGAT 828

QY 81 CySerGlnIleValIleValIleValIleValIleValIleValIleValIleValIleVal 100

DB 829 TCGAGTCAGAAAGTTGATTTACATGATCAAGCTCCCAAGATGTTCTCTCTCTTCA 888

QY 101 AspCyPheLeuThrValIleuGlnIleValIleValIleValIleValIleValIleVal 120

DB 889 GACTGTTTCTCATCTGATCTTCTGGGTAACTGGAAGAGCTTCAACTGTTCACTG 948

QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCyPheProGlyLeuCyGlnIleValIleVal 140

DB 949 CTTCAGAGTGGTTTCTGATGTTCTCTGTTTCTGCTGCTCTGTTGAGAAATCC 1008

QY 141 ThrLeuValProThrCySerIleSerGlnProCySerProValAlaAsnIleGlyProThr 160

DB 1009 ACTCTAGTCCCTACCTGATTTCTGAGCTTGTCTTACCTGTTGCCAATGTTGGCCAACTC 1068

QY 161 ArgIleLeuProAsnLeuTyrIleuGlyCyGlnIleValIleValIleValIleValIle 180

DB 1069 CGAATTTCTCCCAATCTTATCTTGGCTCCAGCGAGATGCTCTCAACAGAGCTGATG 1128

QY 181 GlnGlnIleGlyIleGlyTyrValIleuAsnIleSerTyrThrCyAspProIleAspPhe 200

DB 1129 CAGCAGATGGAGTTGTTATGTTTAAATGCCAGCAAACTCTGCAAGCTGACTTT 1188

QY 201 IleProIleSerIlePheLeuArgValProValAsnAspSerPheCyGlnIleValIle 220

DB 1189 ATCCCGAGATCTCATTTCTGCGGTGCTGGAATGACAGCTTTGTGAGAAATTTTG 1248

QY 221 ProThrLeuAspIleSerValIlePheIleGlnIleValIleValIleValIleValIle 240

DB 1249 CCGTGGTTGAGCAAACTGATGTTTCAATGAGAAAGCAAAAGCTTCAATGATGTT 1308

QY 241 LeuValIleHisCyLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260

DB 1309 CTAGTCACTGTTTACTGGAATCTCCCTCCGACCACTATGCTATACATCAATG 1368

QY 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValIleGlnIleValIleVal 280

DB 1369 AAGAGATGATGATGCTTATGATGAAAGTTTCAAGTTTGTGAAGAAAAAGACTACT 1428

QY 281 IleSerProAsnPheAsnPheLeuGlnIleValIleValIleValIleValIleValIle 300

DB 1429 ATATCTCCAACTTCAATTTTGTGGCCAACTCTGATGAGAAAGATTAAGAAC 1488

QY 301 GlnThr 302

DB 1489 CAGACT 1494

RESULT 14

ABV26680

ID ABV26680 standard; cDNA; 5145 BP.

XX

AC ABV26680;

XX

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 26671.

XX

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 3451; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	3,44e-175	Length:	5145
Score:	1552.00	Matches:	300
Percent Similarity:	99.67%	Conservative:	1
Best Local Similarity:	99.34%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-109_COPY_1_302 (1-302) x ABV20978 (1-5145)

QY 1 MetAlhiSGluMetilegIyThrgInileValThrgIuArgLeuValAlaleuLeuGln 20
DB 589 ATGGCCCATGAGATGATGGAACTCAATGTTACTAGAGGTTGGCTGTGCTGGAA 648
QY 21 SerGIyThrgIuArgLeuValleuValleuSerArgProPheValGluTyraAnthrSer 40
DB 649 AGTGGAAACGAAAGAGCTGCTAATTGATGACCGCATTTGTGAAATACAAATACATCC 708
QY 41 HisIleuGlnIuAlaleuValleuValleuSerArgLeuMetIySArgArgLeuGln 60
DB 709 CACATTTGGAAACCATTAATATCACTGCTCCAGCTTATGAAGCAAGTTGCAACAG 768
QY 61 AspIyValleuIleThrgIuLeuIleGlnHisSerAlaIyShiIyValIleAsp 80
DB 769 GACAAAGTGTAAATACAGAGCTATCCAGCATTCAGCAACATTAAGTTGACATTGAT 828
QY 81 CysSerGlnIyValIleValIleTyraArgIleSerSerGlnAspValAlaSerIleuSer 100
DB 829 TGCAGTCAGAAAGTTGATGATTAAGATCAAGATCCCAAGATGTGGCTCTCTCTTCA 888
QY 101 AspCysPheLeuThrValIleuLeuGlyIyValleuGlyIyValleuSerPheAsnSerVal 120
DB 889 GACTGTTTCTCACTGACTTCTGGGTAACTGGAAGAGCTTCAACTGTTCACCTG 948
QY 121 LeuAlaGlyIyPheAlaGluPheSerArgCysPheProGlyLeuGlyIyIySer 140
DB 949 CTGGCAGGTGGTTGGTGAATTCCTCGTTGTTCCCTGGCTCTGTGAAGAAATCC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 1009 ACTCTAGTCCCTACCTGCAATTCCTGAGCTTCTGCTTACCTGCTTCAACATTTGGCCAAACC 1068
QY 161 ArgIleuProAsnLeuTyraLeuGlyCysGlnArgAspValIleuAsnIyValleuIle 180
DB 1069 CGAATCTTCCCAATCTTATCTTGGTCCAGCGAGATGTCCTCAACAAAGAGCTGATG 1128
QY 181 GlnGlnAsnGlyIleGlyTyraValIleuAsnIleSerTyThrCysProIyProAspPhe 200
DB 1129 CACGCAATGGATGGTATGTTGTTAAATGACAGCAATACCTGTCCAAAGCTGACTTT 1188

QY 201 ILeProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluIyIleLeu 220
DB 1189 ATCCCGAGTCTCATTTCTGCTGTGCTGCTGCAATGACAGCTTTGTGAGAAATTTTG 1248
QY 221 ProTrpLeuAspIySerValAspPheIleGlyIyValIleValIleAsnArgIyCysVal 240
DB 1249 CCGTGTGGACAAATCAGTACATTCATTGAGAAAGCAAAAGCTTCAATGATGTGTT 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 260
DB 1309 CTAGTCACTGTTTACCTGGATCTCCGCTCCGCAACATCCCTATCGCTTACATCATG 1368
QY 261 LysArgMetAspSerLeuAspGluAlaTyraArgPheValIyGluIySArgProThr 280
DB 1369 AACAGATGACATGCTTTTACATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1428
QY 1429 ATATCTCCAAACTTCAATTTTCTGGGCAACCTCGGACTATGAGAAAGATTAAGAAC 1488
DB 301 GlnThr 302
DB 1489 CAGACT 1494

Search completed: June 21, 2004, 21:33:10
Job time : 347.168 sec

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 21:18:06 ; Search time 372.269 Seconds
(without alignments)
3716.370 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMGTQVTRERLVALE.....PNEFGLQGLDYERKIKNGT 302

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3017426 seqs, 2290544650 residues
Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPRO.spool/US10029345/runcat 21062004 164146 29071/app_query.fasta_1.1294
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rmpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=tblomsum62
-TRANS=human4.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINTEN=0
-MAXLEN=200000000 -USER=US10029345 @CGN 1 1 333 @runcat 21062004 164146 29071
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCF_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCFUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length DB	ID	Description
------------	-------------	-----------	----	-------------

1	1552	99.2	1998	9	US-09-816-494-3	Sequence 3, App1
2	1552	99.2	1998	16	US-10-377-072-27	Sequence 27, App1
3	1552	99.2	2071	13	US-10-072-012-257	Sequence 257, App
4	1552	99.2	2732	13	US-10-168-506-2	Sequence 2, App1
5	1552	99.2	3059	17	US-10-257-026-1	Sequence 1, App1
6	1552	99.2	3496	9	US-09-964-277-1	Sequence 1, App1
7	1552	99.2	3544	9	US-09-816-494-1	Sequence 1, App1
8	1552	99.2	3544	16	US-10-377-072-25	Sequence 25, App1
9	1552	99.2	3625	13	US-10-425-114-26234	Sequence 26234, A
10	1552	99.2	3766	13	US-10-343-357-17	Sequence 17, App1
11	1552	99.2	4790	16	US-10-648-593-115	Sequence 115, App
12	1545	98.8	2102	17	US-10-094-749-673	Sequence 673, App
13	1544	98.7	2666	13	US-10-296-115-520	Sequence 520, App
14	1504.5	96.2	2200	13	US-10-072-012-255	Sequence 255, App
15	1194.5	76.4	3332	9	US-09-964-277-20	Sequence 20, App1
16	985	63.0	2453	14	US-10-005-858-1	Sequence 1, App1
17	943	60.3	2476	13	US-10-220-120-28	Sequence 28, App1
18	837	53.5	2039	13	US-10-072-012-265	Sequence 265, App
19	469	30.0	1337	9	US-09-964-899-46	Sequence 46, App
20	469	30.0	1449	12	US-10-052-482-192	Sequence 192, App
21	469	30.0	1830	15	US-10-346-356-1	Sequence 1, App1
22	469	30.0	1909	12	US-10-052-482-191	Sequence 191, App
23	469	30.0	1909	13	US-10-058-270A-129	Sequence 129, App
24	468	29.9	1452	12	US-10-052-482-189	Sequence 189, App
25	468	29.9	1977	12	US-10-052-482-188	Sequence 188, App
26	467.5	29.9	2109	17	US-10-305-720-1135	Sequence 1135, Ap
27	467.5	29.9	2109	16	US-10-641-643-946	Sequence 946, App
28	467.5	29.9	2330	13	US-10-342-887-698	Sequence 698, App
29	467.5	29.9	2330	13	US-10-172-118-698	Sequence 698, App
30	467.5	29.9	2360	15	US-10-175-523-183	Sequence 183, App
31	467.5	29.9	2380	16	US-10-199-221-4	Sequence 4, App1
32	467.5	29.9	2380	17	US-10-117-597-4	Sequence 4, App1
33	467.5	29.9	2649	16	US-10-133-937-48	Sequence 48, App1
34	467.5	29.9	2649	16	US-10-159-563-48	Sequence 48, App1
35	467.5	29.9	2792	10	US-09-971-329-69	Sequence 69, App1
40	461.5	29.5	2104	16	US-10-388-934-654	Sequence 654, App
41	461.5	29.5	2104	16	US-10-191-803-270	Sequence 270, App
42	461	29.5	3286	10	US-09-971-392-9	Sequence 158, App
43	426	27.2	877	9	US-09-764-853-158	Sequence 1100, Ap
44	426	27.2	2240	16	US-10-305-720-1100	Sequence 4, App1
45	426	27.2	2241	16	US-10-174-460-4	

ALIGNMENTS

RESULT 1
US-09-816-494-3
; Sequence 3, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR APPLICATION NUMBER: 2001-03-23
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-816-494-3
Alignment Scores:
Pred. No.: 1.89e-201 Length: 1998

Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%
Query Match: 99.23%
DB: 9

Matches: 300
Conservative: 1
Mismatch: 1
Indels: 0
Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-816-494-3 (1-1998)

```
QY 1 MetAlaHISGLUmetilegIYThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 1 ATGGCCCATGAGATGATTCGAACTCAAAATTGTTACTGAGAGGTGGCTCTCTCGAA 60
QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluIYrAsnThSer 40
DB 61 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCCGCCATTGTGGAAATCAATACATCC 120
QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleuMetIYsArgArgLeuGln 60
DB 121 CACATTTTGGAAAGCATTAATATCACTGCTCCAGCTTATGAAAGCAAGGTTCACAG 180
QY 61 AspIYsValIleuIleThrGluLeuIleGlnHisSerAlaYsHisIYsValAspIleAsp 80
DB 181 GACAAAGTGTAAATTAAGAGCTCATCCAGCATTCACGAAACATAAGTTGACATTGAT 240
QY 81 CysSerGlnYsValIleValIYrAspGlnSerSerGlnAspValAlaSerIleuSer 100
DB 241 TGCAGTCAAGAGGTGTAGTTAGATCAAGATCAAGCTCCCAAGATGTGGCTCTCTCTTCA 300
QY 101 AspCysPheLeuThrValIleuLeuGluYIYsLeuGluYIYsSerPheAsnSerValHisIleu 120
DB 301 GACGTTTCTCACTGTAATCTTGGGTAATCGAAGAGCTTCAACTCTGTTCACTG 360
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlySer 140
DB 361 CTTCAGAGTGGGTTTGTGAGTTCCTCGTTGTTCTCGGCTCGCTGTAAGAAATTC 420
QY 141 ThrIleuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 421 ACTTACTGCTCAACCTGCACTTCTGAGCTTCTGTTACCTGTCGCAACATTTGGCCAAAC 480
QY 161 ArgIleLeuProAsnLeuIYrLeuGlyCysGlnArgAspValLeuAsnIYsGluLeuIle 180
DB 481 CGAATCTTCCCAATCTTATCTTGCTGCCAGCGAGATGCTCTCAACAAGAGCTGATG 540
QY 181 GlnGlnAsnGlyIleGlyIYrValIleuAsnAlaSerIYrThrCysProIYsPheAspPhe 200
DB 541 CAGCAGAAATGGAGATTCGTTATGTGTTAAATGCAAGCAATACCTGTCCAAGCTGACCTT 600
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluYsIleu 220
DB 601 ATCCCGAGACTCATTTCTCGCTGCTGCTGCAAGTACAGCTTTTGGAAATTTTG 660
QY 221 ProTrpLeuAspIYsSerValAspPheIleGlyValAlaYsAlaSerAsnGlyCysVal 240
DB 661 CCGTGTTCGAAATTCAGTATGATTCATTCAGTGAAGCAAAAGCTCCCAATGATGTGTT 720
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIYrIleMet 260
DB 721 CTAGTGCACTGTTAGTGGGATCCCGGCTCCGCAACATCGCTATACATCATG 780
QY 261 LysArgMetAspMetSerIleuAspGluAlaIYrArgPheValIYsGluYsArgProThr 280
DB 781 AAGAGAGTGAAGATCTTTAGATGAACTTACGATTGGAAGAAAGAAAGAGCTTACT 840
QY 841 IleSerProAsnPheAsnPheLeuGlyGlnIleuAsnAspIYrGluYsIYsIleYsAsn 300
DB 841 ATATCTCAAACTCAATTTTCTGGGCCAATCTTGGACTATGAGAAAGATTAAAGAAC 900
QY 301 GlnThr 302
DB 901 CAGACT 906
```

RESULT 2

```
US-10-377-072-27
; Sequence 27, Application US/10377072
; Publication No. US20040009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Mkyoung
; APPLICANT: Tsai, Feng-Ying
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; FILE REFERENCE: MP103-0180NM1M
; CURRENT APPLICATION NUMBER: US/10/377,072
; CURRENT FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1998)
; US-10-377-072-27

Alignment Scores:
Pred. No.: 1,896-201
Score: 1552.00 Length: 1998
Percent Similarity: 99.67% Matches: 300
Best Local Similarity: 99.34% Conservative: 1
Query Match: 99.23% Mismatches: 1
DB: 16 Indels: 0 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x US-10-377-072-27 (1-1998)
QY 1 MetAlaHISGLUmetilegIYThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 1 ATGGCCCATGAGATGATTCGAACTCAAAATTGTTACTGAGAGGTGGCTCTCTCGAA 60
QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluIYrAsnThSer 40
DB 61 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCCGCCATTGTGGAAATCAATACATCC 120
QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleuMetIYsArgArgLeuGln 60
DB 121 CACATTTTGGAAAGCATTAATATCACTGCTCCAGCTTATGAAAGCAAGGTTCACAG 180
```



```
Qy 61 AsplyValleuileThrgluLeuileglnHisSerAlaIysHisIleValIAspIleAsp 80
Db 181 GACAAAGTGTAAATTACAGAGCTCATCCAGATTCCAGCGAAACHTAAGGTGACATTGAT 240
Qy 81 CysSerGlnIlyValValIYraapGlnSerSerGlnAspValAlaSerIleSerSer 100
Db 241 TGCAGTCAGAAAGGTGTAGTTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 300
Qy 101 AspCyPheLeuThrValIleuLeuGlyIlyLeuGlnIlySerPheAnsSerValHisIleu 120
Db 301 GACTGTTTCTCACTGACTTCTGGGTAAACTGGAGAAAGCTTCACTGTTCACCTG 360
Qy 121 LeuAlaGlyIlyPheIleagluPheSerArgCyPheProGlyLeuGlyGluGlyIlySer 140
Db 361 CTTCGAGGTGGTTTCTGAGTTCTCTGCTGTTTCCCTGGCTCTGTGAAGAAATCC 420
Qy 141 ThrLeuValProThrCysIleSerGlnProCysIleuProValAlaAsnIleGlyProThr 160
Db 421 ACTCTAGTCCCTACCTGACCTGCTCAGCCTTGCTTACCTGTTCACATTTGGGCCAAC 480
Qy 161 ArgIleLeuProAsnLeuIlyLeuGlyCysGlnArgAspValIleuAsnIlyGluLeuIle 180
Db 481 CGAATCTCTCCCAATCTTATCTTGGCTGCAGGAGATGTCTCAACAAAGAGCTGATG 540
Qy 181 GlnGlnAsnGlyIleGlyIlyValIleuAsnAlaSerIlyThrCysProIlyProAspPhe 200
Db 541 CAGAGATGGATGGTGTATGTTAAATGCAAGCAATACCTGTCCAAAGCTGACTTT 600
Qy 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIlyIleLeu 220
Db 601 ATCCCGAGCTCATTTCCGCGGTGCTCGTGAATACAGCTTTTGTGAAGAAATTTTG 660
Qy 221 ProIleuAspIlySerValAspPheIleGluIlyAlaIlyAlaSerAsnGlyCysVal 240
Db 661 CCGGTGTGACAAATCATGATGATTTCTATGAGAAACAAAGCTCCCAATGATGTGTT 720
Qy 241 LeuValHisCysLeuAlaGlyIlySerArgSerAlaThrIleAlaIleAlaIlyIleMet 260
Db 721 CTAGTGACGTGTAGTGGGATTCCTCGCTCCGCAACATCGCTATCGCTTACATATG 780
Qy 261 LysArgMetAspMetSerIleAspGluAlaIlyArgPheValIlyGluIlyArgProThr 280
Db 781 AAGAGATGACATGCTTTAGATGAGACTTACAGATTTGTGAAGAAAGAAAGCTACT 840
Qy 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspIlyGlyIlyValIlyIleAsn 300
Db 841 ATATCTCAAACTTCAATTTTCTGGCCCACTCCTGAGATAGAGAAAGATTAAAGAAC 900
Qy 301 GlnThr 302
Db 901 CAGACT 906
```

RESULT 3

```
US-10-072-012-257
; Sequence 257, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernov, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zetserben, Bryan
; APPLICANT: Patursajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esna
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taulier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
```

```
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsdbrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072, 012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265, 102
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265, 514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265, 517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265, 412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265, 395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266, 406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266, 767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267, 057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266, 975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267, 459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 257
LENGTH: 2071
TYPE: DNA
ORGANISM: Homo sapiens
US-10-072-012-257

Alignment Scores:
Pred. No.: 2, 01e-201 Length: 2071
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
Gaps: 0
DB: 13

US-10-029-345A-109_COPY_1_302 (1-302) x US-10-072-012-257 (1-2071)
Qy 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaIleuLeuGlu 20
Db 61 ATGGCCCATGAGATGATGGAGCTCAAAATGTTACTGAGAGGTTGGTGGCTGGCGAA 120
Qy 21 SerGlyThrGluIlyValIleuLeuIleAspSerArgProPheValGluIlyAsnThrSer 40
Db 121 AGTGAACGAAAGAGTGTCTGCTAATTGATGAGCCGCCCATTTGTGAAATCAATACATCC 180
Qy 41 HisIleuGlnAlaIleAsnIleAsnCysSerIlyLeuMetIlyArgArgLeuGln 60
Db 181 CACATTTTGAACCATTAATATATCACTGCTCCAAAGCTTATGAAGCGAAAGTTGCAACAG 240
Qy 61 AsplyValleuileThrgluLeuileglnHisSerAlaIysHisIleValIAspIleAsp 80
Db 241 GACAAAGTGTAAATTACAGAGCTCATCCAGATTCCAGCGAAACHTAAGGTGACATTGAT 300
Qy 81 CysSerGlnIlyValValIYraapGlnSerSerGlnAspValAlaSerIleSerSer 100
Db 301 TGCAGTCAGAAAGGTGTAGTTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 360
Qy 101 AspCyPheLeuThrValIleuLeuGlyIlyLeuGlnIlySerPheAnsSerValHisIleu 120
Db 361 GACTGTTTCTCACTGACTTCTGGGTAAACTGGAGAAAGCTTCACTGTTCACCTG 420
```

```
Alignment Scores:      3.24e-201
Pred. No.:            1552.00
Score:                 99.67%
Percent Similarity:    99.34%
Best Local Similarity: 99.23%
Query Match:
```

Length: 2732
Matches: 300

```
Conservative: 1
Mismatches: 1
Indels: 0
```

QY 1 MetAlaHisGlnIuNetIeGlyThrglnIleVal1ThrglnArgLeuValAlaIeuLeuGln 20
Db 538 ATGGCCCATGAGATGATGGAACTCAATATTGTACTGAGAGCGTTGTGGCTTCTCTGAA 59

QY 21 SerGlyThnGlulysValleuleuileaspserArgProphevalGlutyrAsnthrsr 40
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 598 AGTGGAAACGGAAAAAGTGCGCTAATTGATGACCCGCCCATTTGTGGAATACAATACATCC 65

[illegible]

Dy 61 AspLysValIleuIeThgJluLeuIecInHissSerAlaYshIglyValAspIleasp 80
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 718 GACAAAGTGTATATACAGACTATTCAGATTCAAGAAACAATAAGTTGACATTGAT 777
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 81 CYSserGlnItyrValValTyrAspGlnserGlnspValAlaSerLeuSerSer 100
|||
778 TGCACGTACAGAGGTGTAGTTACATCAAGCTCCCAAGATGTTCCTCTCTCTCA 837
Db

QY 101 AspCysPheLeuThrValLeuGlyValylLeuGluLysSerPheAsnSerValHisLeu 120
DQ 838 GACTCTTTTCACGTACTCTCTGGGAAACTGGAGAAAGACTTTCACCTGTGTTACCTG 897

Qy 121 LeuAlaIyglYpHeaIaGlupHeSerArgCyspHeProGlyLeuCysGluGlyLysSer 140
Db 898 CTTGAGGTGGGTTTGGTGAATTCCTTCGATTGTTCCCTGGCCCTGTGAAGGAAATATCC 957

141 ThrLeuValProThrcysIleSerGlnProCysLeuProValAlaIasnIleGlyProThr 160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971

161 ArgIleuProAsnLeuTyrLeuGlyCyvGlnArgAspValLeuAsnysgIureuIle 180
1018 CGAATTCCTCCCAATCTTTATCTTGGCTGCGACGGAGATGTCCTCAACAGAGAGCTGGT 1027	

181 GlnGlnasndglylleetytyrValleuanaalaserTyrThrCysProlysProasphe 200
|||
1078 CAGAGATGGGATTGGTATGCTTAATGTCAGCAATACCTGTCGAAGGCTGACTTT 1137

201 ILProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlnuylalleu 220
|||||
1138 ATCCCGAGCTCATTTCTCGGTGGCTGTGAATGACAGCTTTTGGAGAAATTTGG 1197
|||||

221 ProTryptophanSerValAspPheIleGluLysAlaLysAspArgGlyCysVal 240
|||
1198 CCTGGTGGCAAAATCACTAGATTTCATGTGAAGAGAAAGACGCCAATGATATGCTTT
|||

241 LeuValHisCysLeuAlaGlyTLeSerArgSerAlaThrIleAlaIleAlaTyrlIleMet 260
1258 CTAAGTCACATGTATTAGCTGGGAACTCCCGCATCCGCACCAATGCGATATAGCTGATAATATATTTT

261 LysArgMetAspMetSerLeuaspGlnAlaTyrArgPheValLysGluLysArgProThr 280
1318 AAGAGATGGACATCTCTTTAGTCAGACTACAGATTGTGCAGAAAAAACCCCTCCTT
1319

Y

281 IIESErProAsnPhaasnPheLcUGlnLeuLeuspyrrGylglsyIleLVasn 300
|||||
|||

b

1378 ATATGTCCAACCTTCATTCTTGGGCCAACTCCCTGA CTTTGAGCAACA CTCTTTTT | |||||
|||||
|||||

	1437
y	301 GlnThr 302	
b		
	1438 CAGACTT 1443	

RESULT 5
S-10-257-026-1
Sequence 1, Application US/10257026
Publication No. US20040086859A1
GENERAL INFORMATION:

```

; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10X.DMS
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3059
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)..(2121)
US-10-257-026-1

Alignment Scores:
Pred. No.: 3,93e-201 Length: 3059
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 17 Gaps: 0

US-10-029-345a-109_copy_1_302 (1-302) x US-10-257-026-1 (1-3059)

Qy 1 MetAlAHieGluMetIlegIYThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
Db 127 ATGGCCCATAGATGATGGAACTCAAAATGTTACTGAGAGGTGGGCTCGCTGGAA 186
Qy 21 SerGIYThrGluYsValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 187 AGTGGAAAGGAAAGAGCTGCTTATGATAGCCGCACTTTGGTAATTCATACATCC 246
Qy 41 HisIleLeuGlnAlaIleAsnIleAsnCySerIleuLeuLeuLeuLeuLeuGln 60
Db 247 CACATTTGGAGCCATTATATCACTGCTCCAACTTATGAAGCAAGTTGCAACAG 306
Qy 61 AspIyValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Db 307 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCAAGCAATAGGTTGACATTGAT 366
Qy 81 CySerGlnIleValIleValIleValIleValIleValIleValIleValIleVal 100
Db 367 TGCAGTCAGAGGTTGATGATTCAGATCAAGCTCCCAAGATGTTGCTCTCTTCA 426
Qy 101 AspCyPheLeuThrValIleuLeuGlyIleuGlnIleuSerPheAsnSerValIleu 120
Db 427 GACTGTTTCTCACTGATCTTGGGTAAGTGGAGAAAGCTTCACTGTTCACTG 486
Qy 121 LeuAlaGlyIlePheAlaGluPheSerArgCyPheProGlyLeuCyGluGlyIleuSer 140
Db 487 CTTCGAGGTGGTGGTTCGATGCTCTGTTGTTCCCTGCTGCTGCAAGGAAATCC 546
Qy 141 ThrLeuValProThrCySIIeserGlnProCyLeuProValAlaAsnIleGlyProThr 160
Db 547 ACTCTAGTCCCTACCTGATTCCTCAGCTTGTACTTCTGCAACATTTGGCAACC 606
Qy 161 ArgIleuProAsnLeuIleuGlyIleuGlnIleuSerPheValIleuAsnIleuLeu 180
Db 607 CGAATCTTCCCATCTTATCTTGGCTGCGAGAGATGCTCTCAACAGAGAGCTGATG 666
Qy 181 GlnGlnAsnGlyIleGlyIleValLeuAsnIleSerThrCyPheProIleuProAspPhe 200
Db 667 CACACAGATGGAGTGTGTATGTGTTAAATGCAAGCATATCTGTCMAAGCTGACTTT 726
Qy 201 IleProGlnSerIlePheLeuArgValProValAsnAspSerPheCyGluIleuLeu 220
Db 727 ATCCCGCAGCTCATTTCTGCGGTGCTGCTGATATACAGCTTTGTGCAAAATTTTG 786
Qy 221 ProTyrPheAspIleuSerValAspPheIleGlyIleuValIleuValIleuValIleu 240
Db 787 CCGTGGTGGCAAAATCATGATGATTCATTGAGAAAGCAAAAGCTTCATGATGTTG 846
```

```

Qy 241 LeuValHisCyPheLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleMet 260
Db 847 CTAGTGACTGTTTATGCTGGATCTCCGCTCCGCAACATTCCTATGCTTACATATG 906
Qy 261 LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValIleuGlyIleuArgProThr 280
Db 907 AAGAGATGGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 966
Qy 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyIleuValIleuAsn 300
Db 967 ATATCTCCAAACTTCAATTTCTGGGCAACTCTGATGATGAGAAAGATTAAGAAC 1026
Qy 301 GlnThr 302
Db 1027 CAGACT 1032

RESULT 6
US-09-964-277-1
; Sequence 1, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125,434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-1

Alignment Scores:
Pred. No.: 4,94e-201 Length: 3496
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 9 Gaps: 0

US-10-029-345a-109_copy_1_302 (1-302) x US-09-964-277-1 (1-3496)

Qy 1 MetAlAHieGluMetIlegIYThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
Db 562 ATGGCCCATAGATGATGGAACTCAAAATGTTACTGAGAGGTGGGCTCGCTGGAA 621
Qy 21 SerGIYThrGluYsValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 622 AGTGGAAAGGAAAGAGCTGCTAATTTGATAGCCGCACTTTGTGAAATTCATACATCC 681
Qy 41 HisIleLeuGlnAlaIleAsnIleAsnCySerIleuLeuLeuLeuLeuLeuGln 60
Db 682 CACATTTGGAGCCATTATATCACTGCTCCAACTTATGAAGCAAGTTGCAACAG 741
Qy 61 AspIyValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Db 742 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAACATAGGTTGACATTGAT 801
Qy 81 CySerGlnIleValIleValIleValIleValIleValIleValIleValIleVal 100
Db 802 TGCAGTCAGAGGTTGATGATTCAGATCAAGCTCCCAAGATGTTGCTCTCTTCA 861
Qy 101 AspCyPheLeuThrValIleuLeuGlyIleuGlnIleuSerPheAsnSerValIleu 120
Db 862 GACTGTTTCTCACTGATCTTGGGTAAGTGGAGAAAGCTTCAACTGTTCACTG 921
Qy 121 LeuAlaGlyIlePheAlaGluPheSerArgCyPheProGlyLeuCyGluGlyIleuSer 140
Db 922 CTTCGAGGTGGTGGTTCGATGTTCTCTGTTGTTTCCCTGCTGCTGTAAGGAAATCC 981
```

QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db ACTGAGTCCCTACTGCACTTCTCAGCTGCTTACTGCTTGTGCAACATTTGGGCAACC 1041
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGluLeuIle 180
Db CGAATTCCTCCCAATCTTATCTTGCTGCGCAGCAGATGTCTTCACACAGAGCTGATG 1101
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
Db CAGCAGAAATGGAGTTAGTTAGTTAATGCAAGCAATACCTGTGTCACAAAGCTGACTTT 1161
QY 201 IleProGlnSerHisPheLeuArgValProValAsnSerPheCysGluLysIleLeu 220
Db ATCCCGAGTCTCATTTCTCGGTGTGCTGTGATGACAGCTTTTGTGAGAAATTTTG 1221
QY 221 ProTrpLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsnGlyCysVal 240
Db CCGTGTGTGACAAATCAGTATGATTTCAATTGAGAAAGCAAGCCTCCAAATGATGTGT 1281
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db CTAGTGAAGTGTGTTAGTGGAGTCTCCGCTCCGACCAATGCTATGCTTACATCATG 1341
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
Db 1342 AAGAGGATGACATGTCTTTAGATGAGACTTACAGATTGTGAAAGAAAAAGACCTACT 1401
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLysIleLysAsn 300
Db 1402 ATATCTCCAAATCTCAATTTCTGTGGCCAACTCCCTGAGCTATGAGAAAGATTAAAGAAC 1461
QY 301 GlnThr 302
Db 1462 CAGACT 1467

RESULT 7

US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589)...(2583)
US-09-816-494-1

Alignment Scores:

Pred. No.: 5,06e-201 Length: 3544
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 9 Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-816-494-1 (1-3544)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
|||||

Db 589 ATGGCCCATGAGATGATTGGAACTCAATTTGTTACTGAGAGGTTGGCTGCTGGA 648
QY 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGlyTyrAsnThrSer 40
Db 649 AGTGGAAAGGAAAAAGTGGCTGCTTATTTGATGAGCGGCAATTTGTGAAATACATAC 708
QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerLysLeuMetLysArgGluGlnGln 60
Db CACATTTGGAAACCAATTAATCAACTGCTCCAAAGCTTATGAAAGAGAGTTGCAACAG 768
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaHisIleValAspIleAsp 80
Db 769 GACAAAGTTAATTAACAGAGCTATCCAGCATTCAGCAAAATTAAGTTGACATTGAT 828
QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 829 TGCAGTGAAGGTTGATGTTAGTACATCAAAAGCTCCAAAGATGTGGCTCTCTCTTCA 888
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
Db 889 GACTGTTTCTCACTGTAATCTTGGGTAAACTGAGAAAGACTTCAACTGTTCACCTG 948
QY 121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
Db 949 CTGAGAGTGGGTTGTGTGAGTCTCTCGTGTTCCTGAGCTCTGTGAGAAATATCC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 1009 ACTGTAGTCCCTACCTGCAATTTCTCAGCCCTGTGTTACTGTTCACAAATTTGGGCCAAC 1068
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGluLeuIle 180
Db 1069 CGAATTCCTCCCAATCTTATCTGTGGCTGCGCAGATGTCTCCAAAGAGAGCTGATG 1128
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
Db 1129 CAGCAGAAATGGAGTTAGTTAGTTAATGCAAGCAATACCTGTGCAAAAGCTGACTTT 1188
QY 201 IleProGlnSerHisPheLeuArgValProValAsnSerPheCysGluLysIleLeu 220
Db 1189 ATCCCGAGTCTCATTTCTCGGTGTGCTGTGATGACAGCTTTTGTGAGAAATTTTG 1248
QY 221 ProTrpLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsnGlyCysVal 240
Db 1249 CCGTGTGTGACAAATCAGTATGATTTCAATTGAGAAAGCAAAAGCCTCCCAATGATGTGT 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet 260
Db 1309 CTAGTGAAGTGTGTTAGTGGAGTCTCCGCTCCGACCAATGCTATGCTTACATCATG 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProThr 280
Db 1369 AAGAGATGACATGTCTTTAGATGAGACTTACAGATTGTGAAAGAAAAAGACCTACT 1428
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLysIleLysAsn 300
Db 1429 ATATCTCCAAATCTCAATTTCTGTGGCCAACTCCCTGAGCTATGAGAAAGATTAAAGAAC 1488
QY 301 GlnThr 302
Db 1489 CAGACT 1494

RESULT 8

US-10-377-072-25
; Sequence 25, Application US/10377072
; Publication No. US20040009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Gluckemann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.

APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Miyoung
TITLE OF INVENTION: Novel 25869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: MPIO3-0180NMIM
CURRENT APPLICATION NUMBER: US/10/377, 072
CURRENT FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: US 09/895, 860
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 370
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/723, 806
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 60/187, 455
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/843, 297
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199, 801
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: US 09/861, 801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205, 508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/816, 494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/815, 419
PRIOR FILING DATE: 2001-03-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 3544
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (589) ... (2586)
US-10-377-072-25

Alignment Scores:
Pred. No.: 5 06e-201 Length: 3544
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conserved: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
Gaps: 0
DB: 16

US-10-029-345a-109_COPY_1_302 (1-302) x US-10-377-072-25 (1-3544)

QY 1 MetAlHieGluNectilegYThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 589 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 648
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 649 AGTGAACGGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
QY 41 HisIleLeuGluValIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB 709 CACATTTGGAAAGCCATTAATCACTGCTCCAGCTTATGAGCGAAGGTTGCAACG 768
QY 61 AspLeuValLeuLeuThrGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 769 GACCAAGCTTATTAATCAAGCTCATCACTCACTCACTCACTCACTCACTCACT 828
QY 81 CysSerGlnIleValValValValValValValValValValValValValValVal 100
DB 829 TCGAGTCAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888
QY 101 AspCysPheLeuThrValLeuLeuGluGlyValLeuGluGlySerPheLeuValHisLeu 120

DB 889 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAGAGCTTCACTGTTCACTG 948
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlyValSer 140
DB 949 CTTGAGGAGGAGGTTGCTGAGTCTCTGCTGTTTCTGAGGCTCTGTAAAGGAAATCC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAlaIleGlyProThr 160
DB 1009 ACTTATGCTTCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1068
QY 161 ArgIleLeuProAsnLeuIleGlyCysGlnArgAspValLeuAsnLeuLeuLeuLeu 180
DB 1069 CCAATTTCTTCCCAATCTTATCTTGGCTCCAGCAGATGCTCTCAACAAGAGCTGATG 1128
QY 181 GlnGlnGlnGlyIleGlyTyrValLeuAlaAlaSerTyrThrCysProIleProAspPhe 200
DB 1129 CAGCAGAAATGGATGGTATGTTATGTTAAATGACAGAAATGCTGTCACAAAGCTGACTTT 1188
QY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleLeu 220
DB 1189 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
QY 221 ProIleLeuAspIleSerValAspPheIleGlyValAlaAlaSerArgGlyCysVal 240
DB 1249 CCGTGGTTGACAAATCACTGATGATTTCACTTGAAGAAAGCAAAAGCTCCATGATGTT 1308
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleIleMet 260
DB 1309 CTAGTGCATGTTTGTGCTGATGATCTCCCTCCCTCCGACCATGCTATGCTCAATCATG 1368
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluValArgProThr 280
DB 1369 AAGAGATGACATGCTTCTTAAATGAGCTTACAGATTGTTGTAAGAAAAAGAACTTACT 1428
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluValIleValAsn 300
DB 1429 ATATCTCCAAATCTTCAATTTTCTGGGCCAACTCTGACATGAGAAAGATTAAAGAAC 1488
QY 301 GlnThr 302
DB 1489 CAGACT 1494

RESULT 9

US-10-425-114-26234

/ Sequence 26234, Application US/10425114

/ Publication No. US20040034888A1

/ GENERAL INFORMATION:

/ APPLICANT: Liu, Jindong

/ APPLICANT: Zhou, Yihua

/ APPLICANT: Kovacic, David K.

/ APPLICANT: Screen, Steven E.

/ APPLICANT: Tabaska, Jack E.

/ APPLICANT: Cao, Yongwei

/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

/ FILE REFERENCE: 38-21(53313)B

/ CURRENT APPLICATION NUMBER: US/10/425, 114

/ NUMBER OF SEQ ID NOS: 73128

/ SEQ ID NO 26234

/ LENGTH: 3625

/ TYPE: DNA

/ ORGANISM: Homo sapiens

/ FEATURE:

/ OTHER INFORMATION: Clone ID: LIB4119-028-H6_FU1

US-10-425-114-26234

Alignment Scores:
Pred. No.: 5.26e-201 Length: 3625
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conserved: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0

DB: 13 Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x US-10-425-114-26234 (1-3625)
QY 1 MetAlahisgluMecIleglyThrglnlLevalThrgluArgleuValAlaleuLeuglu 20
Db 692 ATGGCCCATGAGATGATGGAATCAAAATGTTACTGAGAGGTTGGTGGCTCTGCGAA 751
QY 21 SerGlyThrgluysValleuLeuileasPserArgProPheValGluTyraAsnThrSer 40
Db 752 AGTGAACGGAAGAAAGCTGCTATATGATGAGCCGCGCATTTGTGGAATCAATAACATCC 811
QY 41 HisIleuenglualaleaenlleasnCyserIysleuMecIysArgArgLeugln 60
Db 812 CACATTTGGAAAGCATTAATATCACTGCTCCAGCTTAAGAACGGAAGTTGCCAAG 871
QY 61 AspIysValleuilethrgluLeuileglnhisSerAlaYshIlyValAspIleasp 80
Db 872 GACAAAGTTAAATTACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 931
QY 81 CysSerGlnIysValValIyTyraSpGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 932 TGCAGTCAGAAAGTTGATGATTACGATCAAGCTCCAAAGATGTGCTCTCTCTTCA 991
QY 101 AspCysPheLeuThrValleuLeuglyIySleugIyIysSerPheAsnSerValHisleu 120
Db 992 GACTGTTTCTCACTGATCTTCTGGGTAAATGGAAGAAAGCTTCACTCTTCACTG 1051
QY 121 LeuAlaglyglyPheAlagluPheSerArgCysPheProglyLeuGlyIySser 140
Db 1052 CTTCAGAGTGGGTTGGTGAAGTCTCTGTTGTTCCCTGCGCTCTGGAAGGAAGATCC 1111
QY 141 ThrIleuValProthrgIysIleSerGlnProCysleuProValAlaAsnIleGlyProthr 160
Db 1112 ACTGTAGTCCCTACCTGATTTCTGAGCTTGTCTTACTGTTGCCAACATTTGGGCAACC 1171
QY 161 ArgIleuProAsnLeuTyrlleuglyCysgluArgAspValleuAsnIySgluLeuile 180
Db 1172 CGAATTTCTCCCATCTTATCTTGGCTGCCAGGAGATGTCTTCAACAAAGAGCTGATG 1231
QY 181 GlnGlnAsnGlyIleglyTyValleuAsnAlaSerTyThrgIyProIySProAspPhe 200
Db 1232 CAGAGATGGAGATGGTATGTATTAATCCAGCAATACCTGTCCAAAGCTGACTTT 1291
QY 201 IleProgluSerHisPheLeuArgValProValAsnAspSerPheCysGluIySleu 220
Db 1292 ATCCCCAGTCTCATTTCTGCTGCTGCTGCTGATGACACCTTTTGTGAAGAAATTTTG 1351
QY 221 ProTrpLeuAspIySserValAspPheIlegluysAlaIySAsnGlyCysVal 240
Db 1352 CCGGTGTTGCAAAATGATGATTTCTTGAAGAAAGCAAGCTCCAAATGATGTGTT 1411
QY 241 LeuValHisCysleuAlaglyIleSerArgSerAlaThrIleAlaIleAlaTyrlleuMec 260
Db 1412 CTAATGACATCTTTAGCTGGGATCTCCGCTCCGCCACATCGCTATCGCTTACATCATG 1471
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyraPheValIySgluIySArgProthr 280
Db 1472 AAGGAGATGACATGTTTATGATGAACCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1511
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnleuLeuAspTyrgluIySlyIleLysAsn 300
Db 1532 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTCGAGCTATGAGAAAGATTAAGAAC 1591
QY 301 GlnThr 302
Db 1592 CAGACT 1597
RESULT 10
US-10-343-357-17
; Sequence 17, Application US/10343357
; Publication No. US20040058341A1
; GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.; TANG, Y.Tom
APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
APPLICANT: YAO, Monique G.; BURROD, Neil
APPLICANT: WANG, Yumei B.; STEWART, Elizabeth A.
APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
APPLICANT: LEE, Ernestine A.; HAFALIA, April J.A.
APPLICANT: LU, Dying Anna M.; TRIBOULEY, Catherine M.
APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Marian R.
APPLICANT: YUE, Henry; WARREN, Bridget A.
APPLICANT: NGUYEN, Daniel B.; CHAWLA, Narinder K.
APPLICANT: KEARNEY, Liam
TITLE OF INVENTION: PROTEIN PHOSPHATASES
FILE REFERENCE: PI-0173 PCT
CURRENT APPLICATION NUMBER: US/10/343,357
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: PCT/US01/23716
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,679
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/223,272
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/224,309
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 60/226,728
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 60/229,254
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: US 60/231,366
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 17
LENGTH: 3766
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CBI
US-10-343-357-17
Alignment Scores:
Pred. No.: 5,61e-201 Length: 3766
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
Gaps: 0
US-10-029-345A-109_COPY_1_302 (1-302) x US-10-343-357-17 (1-3766)
QY 1 MetAlahisgluMecIleglyThrglnlLevalThrgluArgleuValAlaleuLeuglu 20
Db 538 ATGGCCCATGAGATGATGGAATCAAAATGTTACTGAGAGGTTGGTGGCTCTGCGAA 597
QY 21 SerGlyThrgluysValleuLeuileasPserArgProPheValGluTyraAsnThrSer 40
Db 598 AGTGAACGGAAGAAAGCTGCTATATGATGAGCCGCGCATTTGTGGAATCAATAACATCC 657
QY 41 HisIleuenglualaleaenlleasnCyserIySleuMecIysArgArgLeugln 60
Db 658 CACATTTGGAAAGCATTAATATCACTGCTCCAGCTTAAGAACGGAAGTTGCCAAG 717
QY 61 AspIysValleuilethrgluLeuileglnhisSerAlaYshIlyValAspIleasp 80
Db 718 GACAAAGTTAAATTACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 777
QY 81 CysSerGlnIysValValIyTyraSpGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 778 TGCAGTCAGAAAGTTGATGATTACGATCAAGCTCCAAAGATGTGCTCTCTCTTCA 837
QY 101 AspCysPheLeuThrValleuLeuglyIySleugIyIysSerPheAsnSerValHisleu 120
Db 838 GACTGTTTCTCACTGATCTTCTGGGTAAATGGAAGAAAGCTTCACTCTGTTCACTG 897

QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlySer 140
| | | | |
Db 898 CTTGAGAGTGGGTTTCTGAGTTCTCTGTTGTTTCCCTGCTGTGAAGAAATCC 957
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAenIleGlyProThr 160
| | | | |
Db 958 ACTCTAGTCCCTACCTGATTTCTGAGCCTTGCTTACTGTTGCAATGAGCCAAACC 1017
QY 161 ArgIleLeuProAsnLeuValLeuGlyCysGlnArgAspValLeuAsnIleGlyLeu 180
| | | | |
Db 1018 CGAATCTTCCCAATCTTATCTTGCTGACGAGATGCTCTCAACAGAGCTGATG 1077
QY 181 GlnGlnAsnGlyIleGlyValLeuAsnAlaSerTyThrCysProIleProAspPhe 200
| | | | |
Db 1078 CAGCAAGATGGATTTGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 1137
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlyIleLeu 220
| | | | |
Db 1138 ATCCCGAGTCTCATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAAAAATTTTG 1197
QY 221 ProTrpLeuAspIleSerValAspPheIleGlyLeuAlaIleAsnIleGlyVal 240
| | | | |
Db 1198 CCGTGGTTGGCAAAATCAGTAGATTTCTATTGAGAAAGCAAAAGCTTCAATGATGTGT 1257
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 260
| | | | |
Db 1258 CTAGTCACTGTTTACTGGATCTCCGCTCCGCAACCTCCCTATCCCTTACATCATG 1317
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyArgPheValIleGlyIleAspProThr 280
| | | | |
Db 1318 AAGAGATGAGCATGCTTTAGATGAAGCTTACAGATTGTGAAGAAAGAAAGACCTACT 1377
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrlleIleValIleVal 300
| | | | |
Db 1378 ATATCTCCAAACTTCAATTTCTGGGCACTCTGAGATATGAGAAAGATTAAAGAAC 1437
QY 301 GlnThr 302
| | | | |
Db 1438 CAGACT 1443

RESULT 11
US-10-648-593-115
; Sequence 115, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648, 593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 115
; LENGTH: 4790
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-648-593-115

Alignment Scores:
Pred. No.: 8,47e-201
Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%
Query Match: 99.23%
DB: 17
Length: 4790
Matches: 300
Conservative: 1
Mismatch: 1
Indels: 0
Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x US-10-648-593-115 (1-4790)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
| | | | |
Db 184 ATGGCCCATGAGATGATTTGGAATCTCAATTTTACTGAGAGGTTGGTCTGCTGGAA 243
QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyraenThrSer 40
| | | | |
Db 244 AGTGAACCGGAAAGAAAGTCTGCTTAATGATAGCCGCCCAATTTGTGGAATACATAC 303
QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerIleMetIleValArgLeuGln 60
| | | | |
Db 304 CACATTTGGAAGCCATTAATATCACTCTCCAAAGCTTATGAAAGGAGTTGCAACAG 363
QY 61 AspIleValIleIleThrGluLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
| | | | |
Db 364 GACAAAGTTTAAATTCAGAGCTTATCCAGCATTCAGCCAAACATTAAGTTGATGAT 423
QY 81 CysSerGlnIleValValIleValIleAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
| | | | |
Db 424 TGCAGTCAGAGGTTTATGATTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 483
QY 101 AspCysPheLeuThrValLeuLeuGlyIleValLeuGlyIleSerPheAsnSerValIleLeu 120
| | | | |
Db 484 GACTGTTTCTCAGTGTACTTCTGGGTAACTGGAGAAAGCTTCAACTCTGTCACTG 543
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlySer 140
| | | | |
Db 544 CTTGAGAGTGGGTTTCTGAGTTCTCTGTTTCCCTGCTGTGAAGAAATCC 603
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAenIleGlyProThr 160
| | | | |
Db 604 ACTCTAGTCCCTACCTGATTTCTGAGCCTTGCTTACTGTTGCCAAATCTGGGCCAAACC 663
QY 161 ArgIleLeuProAsnLeuValLeuGlyCysGlnArgAspValLeuAsnIleGlyLeu 180
| | | | |
Db 664 CGAATCTTCCCAATCTTATCTTGCTGCTGCGAGAGATGCTCTCAACAGAGCTGATG 723
QY 181 GlnGlnAsnGlyIleGlyValLeuAsnAlaSerTyThrCysProIleProAspPhe 200
| | | | |
Db 724 CAGCAAGATGGATTTGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 783
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlyIleLeu 220
| | | | |
Db 784 ATCCCGAGTCTCATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAAAAATTTTG 843
QY 221 ProTrpLeuAspIleSerValAspPheIleGlyValAlaIleAsnIleGlyVal 240
| | | | |
Db 844 CCGTGGTTGGCAAAATCAGTAGATTTCTATTGAGAAAGCAAAAGCTTCAATGATGTGT 903
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 260
| | | | |
Db 904 CTAGTCACTGTTTACTGGATCTCCGCTCCGCAACATGCTATCCCTTACATCATG 963
QY 261 LysArgMetAspMetSerLeuAspGluAlaTyArgPheValIleGlyIleAspProThr 280
| | | | |
Db 964 AAGAGATGAGCATGCTTTAGATGAAGCTTACAGATTGTGAAGAAAGAAAGACCTACT 1023
QY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrlleIleValIleVal 300
| | | | |
Db 1024 ATATCTCCAAACTTCAATTTCTGGGCACTCTGAGATATGAGAAAGATTAAAGAAC 1083
QY 301 GlnThr 302
| | | | |
Db 1084 CAGACT 1089

RESULT 12
US-10-094-749-673
; Sequence 673, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAL, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI


```

APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHITO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 673
LENGTH: 2102
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-673

```

```

Alignment Scores:
Pred. No.: 1,88e-200 Length: 2102
Score: 1545.00 Matches: 299
Percent Similarity: 99.34% Conservative: 1
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 98.79% Indels: 0
Gaps: 0

```

US-10-029-345a-109_copy_1_302 (1-302) x US-10-094-749-673 (1-2102)

```

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 56 ATGGCCCATGAGATGATTGGAACCAATTTGTTACTGAGAGTGGTGGCTCTGCTGAAA 115
QY 21 SerGlyThrGluValLeuLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
DB 116 AGTGAACGGAAGAAAGTCTGCTGTAATGATAGCCGCCCATTTTGAGATACATACATCC 175
QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIleuMetIleValArgLeuGln 60
DB 176 CACATTTTGAAGCATTAATATCACTGCTCAAGCTTATGAGCGAAGGTTGCAACAG 235
QY 61 AspIleValLeuIleThrGlnLeuIleGlnHisSerAlaValHisIleValAspIleAsp 80
DB 236 GACAAAGGTTAATTAACAGAGCTCATCCAGCATTCAGGAAACATPAAGGTTGACATGAT 295
QY 81 CysSerGlnValValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 296 TGCACTCGAAGAGTTGATGATTACATCAAGACTCCAAAGATGTTGCTCTCTCTTCA 355
QY 101 AspCysPheLeuThrValLeuLeuGluIleValLeuGluIleValSerPheAsnSerValHisLeu 120
DB 356 GACTGTTTCTCACTGTAATCTCTGGGTAAACTGAGGAAGAGCTTCACTGTTTCACTG 415
QY 121 LeuAlaGlyValPheAlaGluPheSerArgCysPheProGlyLeuCysGluGluValSer 140
DB 416 CTGGAGAGGAGGTTGCTGAGAGTCTCTGTTTCTCTGCTGCTGAGAGGAAATCC 475
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 476 ACTCTAGTCCCTCACTGATTTCTCAAGCTTCTTACCTGTTGCCAATGTTGGGCAAC 535
QY 161 ArgIleLeuProAsnLeuTyrIleGluCysGlnArgAspValLeuAsnIleGluLeuIle 180

```

```

DB 536 CTAATCTTCCCATCTTTATCTTGGCTGCAGCGAGATGCTCTCAACAAGAGCTGATG 595
QY 181 GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAspPhe 200
DB 596 CAGCAGATGGAGATGGTATGTTATGTTAATGCAAGCAATACCTGCTCAAGCTGACTTT 655
QY 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
DB 656 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715
QY 221 ProThrLeuAspIleSerValIleAspPheIleGluIleValAlaValAlaSerAsnGlyCysVal 240
DB 716 CCGTGGTTGACAAATCAAGATTTCTGAGAAAGCAAAAGCTTCAATGATGTTGTT 775
QY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIle 260
DB 776 CTATGTCACCTGTTTATGCTGGATCTCCCGCTCCGCCACATTCCTATCCCTCATCATG 835
QY 261 LysArgMetAspIleSerLeuAspGluAlaTyrArgPheValIleGluIleValArgProThr 280
DB 836 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACCTACT 895
QY 281 IleSerProAsnPheAsnPheLeuGluIleLeuAspTyrGluIleValIleValAsn 300
DB 896 ATATCTCAAACTTCAATTTCTGGGCCAATCTCTGACATAGAGAAAGATTAGAAC 955
QY 301 GlnThr 302
DB 956 CACACT 961

```

```

RESULT 13
US-10-296-115-520
Sequence 520, Application US/10296115
Publication No. US20040053248A1
GENERAL INFORMATION:
APPLICANT: Hyeq Inc
TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides
FILE REFERENCE: 784PCT
CURRENT APPLICATION NUMBER: US/10/296,115
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: US09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552,317
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 1478
SEQ ID NO 520
LENGTH: 2966
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(2966)
OTHER INFORMATION: n = a,t,c or g
US-10-296-115-520

```

```

Alignment Scores:
Pred. No.: 4.66e-200 Length: 2966
Score: 1544.00 Matches: 299
Percent Similarity: 99.34% Conservative: 1
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 98.72% Indels: 0
Gaps: 0
DB: 13

```

US-10-029-345a-109_copy_1_302 (1-302) x US-10-296-115-520 (1-2966)

```

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 23 ATGGCCCATGAGATGATTGGAACCAATTTGTTACTGAGAGTGGTGGCTCTGCTGAAA 82
QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
DB 83 AGTGAACGGAAGAAAGTCTGCTGTAATGATAGCCGCCCATTTTGAGATACATACATCC 142

```

[illegible]

```

APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsdbrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Removing Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 255
LENGTH: 2200
TYPE: DNA
ORGANISM: Homo sapiens
US-10-072-012-255

Alignment Scores:
Pred. No.: 7,28e-195 Length: 2200
Score: 1504.50 Matches: 300
Percent Similarity: 94.95% Conservative: 1
Best Local Similarity: 94.64% Mismatches: 1
Query Match: 96.20% Indels: 15
DB: 13 Gaps: 4

US-10-029-345A-109_COPY_1_302 (1-302) x US-10-072-012-255 (1-2200)
QY 1 MetAlhIsgIuMeTlIegIYThrgInIleValIhrgIuArygleuValAlaleuIeuglu 20
Db 61 ATGGCCCATGAGTATGATTTGGAATCTCAATTTGTTACTGAGAGGTTGGTGGCTGCTGGAA 120
QY 21 SerGIYThrgIuLyValleuIeulIeapSerAryProPheValGIuTYrAsnThSer 40
Db 121 AGTGGAAACGGAAMAACTGCTGCTAAATTGATACCCGCGCAATTGGGAATTACAAATACATCC 180
QY 41 HIsIILeugIuAlaIleAnIleAnCySerIySleuMeTyAryAryIeugInIln 60
Db 181 CACATTTTGGAAACCTTAATATCAATGCTCTCCAAAGCTTTAGAACGGAAGTTGCAACG 240
QY 61 AspIyValIeulIeThrgIuIeulIeGIhIIsEralAlybHIsIlyS----- 76
Db 241 GACAAAGTGTTATTACAGAGCTCATCCGACATTCAGCGAAACATTAAGTTAAACGCTCAG 300
QY 77 ValIapIleapCySerGIuLyValValValIYrAapGIuSerSerGIuAapValAla 96

```

```

Db      301 GTTGACATGATTGCAGTACAGAGTTGATTACATCAAGCTCCCAAGATGTTGCC 360
Qy      97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsn 116
Db      361 TCTCTCTCTCCAGACTGTTTTCTCAGTGTATCTTGAGTAACCGAGGAAGAGCTTCAAC 420
Qy      117 SerValHisLeuLeuAla-----GlyGlyPheAlaGluPheSerArgCysPhe 132
Db      421 TCTGTACACTGCTTGACAGGTTTATCTTACGTGGGTTTCTGAGTTCTCTCGTTGTTTC 480
Qy      133 ProGlyLeuCysGluGlyLysSerThrLeuValProThrCysIleSerGlnProCysLeu 152
Db      481 CTTGGCCCTGTGTGAAGAAATTCACCTCTAGCTCCCTACTGCACTTCTCAGCCTTGCTTA 540
Qy      153 ProValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyrLeuGlyCysGlnArg 172
Db      541 CTTGTGGCAACATTTGGGCGCAACCGAATCTTCCCAATCTTATCTTGGCTGCAGAGA 600
Qy      173 AspValLeuAsnLysGluLeuIleGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSer 192
Db      601 GATGCTCTCAACAAAGAGACTGATGACAGCAATGGAGTTGTTATGTGTTAAATCCAGC 660
Qy      193 TyrThrCysProLysProAspPheIleProGlnSerHisPheLeuArgValProValAsn 212
Db      661 AATACCTGTCCAAAGCCGTGACTTTATCCCGAGTCTCATTTCTGCGGTGCTGTGAT 720
Qy      213 AspSerPheCysGluLysIleLeuProTyrLeuAspLysSerValAspPheIle----- 230
Db      721 GACAGCTTTGTGTGAGAAATTTTGGCGGTGAGCAATCAGATGATTCAATGGTAAAG 780
Qy      231 -----GluLysAlaLysAlaSerAsnGlyCysValLeuValHisCysLeuAla 246
Db      781 TTGACTTATACAGAGAAAGCAAAAGCTTCCATGATGTGTTAGTGCACATGTTACT 840
Qy      247 GlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLysArgMetAspMetSer 266
Db      841 GGGATCTCCCGCTCGGCACCATCGCTATCGCTCATCATGAAAGAGATGAGACATGTCT 900
Qy      267 LeuAspGluAlaTyr---ArgPheValLysGluLysArgProThrIleSerProAsnPhe 285
Db      901 TTAGTGAAGCTTTCAGAGAGATTGTGTGAAGAAAAAAGCACTACTATATCTCCAAATTC 960
Qy      286 AsnPheLeuGlyGlnLeuLeuAspTyrGluLysLysIleLysAsnGlnThr 302
Db      961 AATTTTCTGGGCCAATCTCTGACTATGAGAAAGATTAAAGAACCAACT 1011

RESULT 15
US-09-964-277-20
; Sequence 20, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 3332
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-964-277-20

Alignment Scores:
Pred. No.: 4,75e-152 Length: 3332
Score: 1194.50 Matches: 244
Percent Similarity: 81.19% Conservative: 2
Best Local Similarity: 80.53% Mismatches: 1
Query Match: 76.37% Indels: 56
DB: 9 Gaps: 1

```

```

US-10-029-345a-109_copy_1_302 (1-302) x US-09-964-277-20 (1-3332)
Qy      1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
Db      562 ATGGCCATAGATGATTGGAACCTCAATTTGTACTGAGAGGTTGGTGCTGTGCGAA 621
Qy      21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db      622 AGTGAAGCAAGAAAGAGTGTGTAATTGATAGCCGGCCATTTGTGGAATCAATACATCC 681
Qy      41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgLeuGln 60
Db      682 CACATTTTGAACCATTAATATCACTGCTCAAGCTTATGAAAGCAAGGTTGCAACAG 741
Qy      61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db      742 GACAAAGTTAATTACAGAGCTCATCCAGCATTCAGCGAAACATGAAGTTGACATTGAT 801
Qy      81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db      802 TGCAAGTCAGAAAGTTGATGATTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 861
Qy      101 AspCysPheLeuThrValLeuLeuGlyLysSerPheAsnSerValHisLeu 120
Db      862 GACTGTTTCTCACTGTACTTCTGGTAACTGAGAAAGCTTCAACTCTGTTCACTTG 921
Qy      121 -LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyLys 140
Db      922 CTTG----- 925
Qy      140 rThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProth 160
Db      925 ----- 925
Qy      160 rArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuI 180
Db      926 -----CAGAGCTGAT 936
Qy      180 eGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAsp 200
Db      937 GCACAGAAATGGAGTTGTATGTAAATGCCAGAAATACGTGTCCAAAGCCTGACTT 996
Qy      200 eIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle 220
Db      997 TATCCCGAGTCTCATTTCTGCGCTGTGCTGTAAGACAGCTTTGTGAGAAATTTT 1056
Qy      220 uProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVa 240
Db      1057 GCCGTGTGTGACAAATCAGATGATTCATTTGAGAAAGCAAAAGCTCCATGATGATGT 1116
Qy      240 IleuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 260
Db      1117 TCTAGTGCACGTGTTAGCTGGGATCTCCCGCTCGCCACCATCGCTATCGCTACATCAT 1176
Qy      260 tLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProth 280
Db      1177 GAAGAGATGAGACATGCTTATGATGAAGCTTACAGATTGTGAGAAAAAAGCACTCAC 1236
Qy      280 rIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLysIleLysAs 300
Db      1237 TATATCTCCAAATTCATATTTCTGGGCCAATCTCTGACTATGAGAAAGATTAAAGAA 1296
Qy      300 nGlnThr 302
Db      1297 CCAGACT 1303

```

Search completed: June 22, 2004, 03:06:44
Job time : 381.269 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2004, 21:10:41 ; Search time 59.9628 Seconds
(without alignments)
2794.987 Million cell updates/sec

Title: US-10-029-345A-109_COPY_1_302
Perfect score: 1564
Sequence: 1 MAHEMIGTIVTERVALLE.....PNEFLGQLDYBKIKKQF 302

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-O=/cgn2_1/USPTO.spool/US10029345/runcat 21062004.164146.29052/app_query.fasta_1.1294
-DB=Issued Patents NA -OPMT=fastaop -SUFFIX=rml -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345 @CEN 1.1 76 @runcat 21062004.164146.29052 -NCPU=6 -ICPU=3
-NO MMP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCFUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1552	99.2	1998	4	US-09-816-494-3	Sequence 3, App1
2	1552	99.2	3544	4	US-09-816-494-1	Sequence 1, App1
3	988	63.2	2377	4	US-09-920-668-3	Sequence 3, App1
4	469	30.0	1830	4	US-09-557-921-1	Sequence 1, App1
5	467.5	29.9	2109	4	US-09-016-434-1135	Sequence 1135, App
6	467.5	29.9	2109	4	US-09-023-655-946	Sequence 946, App
7	426	27.2	2240	4	US-09-016-434-1100	Sequence 1100, App
8	425	27.2	944	4	US-09-371-671B-10	Sequence 10, App1
9	425	27.2	1238	4	US-08-530-290-11	Sequence 11, App1
10	425	27.2	1238	4	US-09-702-705-803	Sequence 803, App
11	425	27.2	1238	4	US-09-736-457-803	Sequence 803, App
12	425	27.2	1238	4	US-09-614-124B-803	Sequence 803, App

13	425	27.2	1238	4	US-09-671-325-803	Sequence 803, App
14	425	27.2	1238	4	US-09-589-184-803	Sequence 803, App
15	425	27.2	2064	4	US-09-702-705-825	Sequence 825, App
16	425	27.2	2064	4	US-09-736-457-825	Sequence 825, App
17	425	27.2	2064	4	US-09-614-124B-825	Sequence 825, App
18	425	27.2	2064	4	US-09-671-325-825	Sequence 825, App
19	425	27.2	2064	4	US-09-589-184-825	Sequence 825, App
20	425	27.2	2109	4	US-09-702-705-826	Sequence 826, App
21	425	27.2	2109	4	US-09-736-457-826	Sequence 826, App
22	425	27.2	2109	4	US-09-614-124B-826	Sequence 826, App
23	425	27.2	2109	4	US-09-671-325-826	Sequence 826, App
24	425	27.2	2109	4	US-09-589-184-826	Sequence 826, App
25	423.5	27.1	1393	2	US-08-990-379-2	Sequence 2, App1
26	415	26.5	2000	4	US-09-016-434-1291	Sequence 1291, App
27	408	26.1	2303	4	US-09-922-146-3	Sequence 3, App1
28	382	24.4	1287	2	US-08-990-379-1	Sequence 1, App1
29	360	24.3	279	4	US-09-016-434-91	Sequence 51, App1
30	357	22.8	1619	4	US-09-702-705-801	Sequence 801, App
31	357	22.8	1619	4	US-09-736-457-801	Sequence 801, App
32	357	22.8	1619	4	US-09-614-124B-801	Sequence 801, App
33	357	22.8	1619	4	US-09-671-325-801	Sequence 801, App
34	357	22.8	1619	4	US-09-589-184-801	Sequence 801, App
35	357	22.8	1619	4	US-09-702-705-804	Sequence 804, App
36	357	22.8	1637	4	US-09-736-457-804	Sequence 804, App
37	357	22.8	1637	4	US-09-614-124B-804	Sequence 804, App
38	357	22.8	1637	4	US-09-671-325-804	Sequence 804, App
39	357	22.8	4637	4	US-09-589-184-804	Sequence 804, App
40	295.5	18.9	539	4	US-09-389-681-311	Sequence 311, App
41	295.5	18.9	539	4	US-09-620-405B-311	Sequence 311, App
42	295.5	18.9	539	4	US-09-339-338-311	Sequence 311, App
43	295.5	18.9	539	4	US-09-433-826B-311	Sequence 311, App
44	295.5	18.9	539	4	US-09-604-287A-311	Sequence 311, App
45	295.5	18.9	539	4	US-09-834-759-311	Sequence 311, App

ALIGNMENTS

RESULT 1
US-09-816-494-3
; Sequence 3, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-816-494-3

Alignment Scores:
Pred. No.: 8.04e-195
Score: 1552.00
Percent Similarity: 99.67%
Best Local Similarity: 99.34%
Query Match: 99.23%
DB: 4
Gaps: 0

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-816-494-3 (1-1998)

QY 1 MetAlHIGlUmetllegYThrgInllevaThhGluArgdeuValAlaLeuGlu 20
DB 1 ATGGCCACGATGATGATGAACTCAATTGTACTGAGAGCTGGTGGCTGGGAA 60
QY 21 SerGlyThrcGluYuvValIleuLeuIleAapSerArgProPhaValGluYrAenThrSer 40

```
Db 61 AGTGAACGAAAAAGCTCTTAATTGATGCGGCGCATTTGGGAATCAATACATCC 120
QY 41 H1S1LEUAGLUALIIEASNIEASNCYSERLYSEUWELYSARGALEUINGLN 60
Db 121 CACATTTGGAAAGCCATTAATATCACTGCTCCAAAGCTTTAGAGCAAGGTTGCAACG 180
QY 61 ASPLYVALLEUILETHRGULEUILEG1NHISERIALYSHIELYVALASP11LEAP 80
Db 181 GACAAAGTGTTAATTAACAAGCTTCATCAGCATTCAGCAACATTAAGTTGACATTTGAT 240
QY 81 CYSERGLNLYSVALVALTYRASPGLNSERSEGLNLYSVALIASERLEUSER 100
Db 241 TGCAGTCAGAAAGTTGATTAGATTACATCAAGCTCCCAAGATTTGCTCTCTCTTGA 300
QY 101 ASPCYSPHELEUHRVALLEUWELGLYLYSEUGLULYSERPHENASERVALHISLEU 120
Db 301 GACTGTTTCTCATCTGACTTCTGGGTMAACTGGAGAGGCTTCAACTCTGTTCACTG 360
QY 121 LEUVALAGLYG1PHEALAGUPHESERARGYSPHPROGLYLEUCYSG1UGLYLYSSER 140
Db 361 CTTCAGAGTGGGTTTGTGAGTTCTCTCGTTTTCCTGGCTCTGTGAGAGAAATCC 420
QY 141 THRLEUVALPROTHRYSIIESERGLNPROCYSEUPROVALIAASNIILEG1YPROTHR 160
Db 421 ACTCTAGTCCCTACCTGCATTTCTCAGCCCTTGCTACTGCTGCCAACATTTGGCCAAAC 480
QY 161 ARG1LEUAPROASNLEUTYRLEUGLYCYSGINARGASPVALLLEUASGLYLEUILE 180
Db 481 CGAATCTTCCCAATCTTTATCTTGAGCTGCGACGAGATGCTCTCAACAGAGGCTGATG 540
QY 181 GING1NASGLY1LEG1YTYRVALLEUASNAIASERYTHRCPROLYSPROASPPHE 200
Db 541 CAGCAAAATGGGATGTGATTATGTTAATGCCAGCAATACCTGTCAAAGCTTACCTT 600
QY 201 ILEPROGLNSERHISPHLEUARGVALPROVALIASNASERPHCYSG1ULYS1LEU 220
Db 601 ATCCCGAATCTCATTTCTCGCGTGCTGTGAATGACAGCTTTTGAGAGAAATTTTG 660
QY 221 PROTRPLEUASPLYSESERVALASPHLEI1EG1ULYSALYSALASERASNG1CYSPVA 240
Db 661 CCGTGTTGGAACAATCAGTATGATTTCATTGAGAAAGCAAGCTCCAAATGATGCTT 720
QY 241 LEUVALHISCYSEUVALAG1Y1LESEARGSERIALATHRIEALIEALTYRI1EWET 260
Db 721 CTAGTGCACTGTTTATGCTGGATCTCCGCTCCGACCATCGCTATCGCTACATCATG 780
QY 261 LYSATGMEIASPMETSERLEUASPG1UALATYRARGPHEVALYSG1ULYVARGPROTHR 280
Db 781 AAGAGATGACATGCTTTAGATGACCTTACAGATTGTTGAAAGAAAGAAAGCTTACT 840
QY 281 ILESERPROASNPHENASPHLEUGLYCINLEUWELYS1YRGLYLYS1LEYSASN 300
Db 841 ATATCTCCAAACTCAATTTCTGGGCCAAGCTCCGACATGAGATGAGAAAGATTAAAGAC 900
QY 301 G1NTHR 302
Db 901 CAGACT 906
```

RESULT 2

```
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 3892 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: PHOSPHATASE MOLECULES AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; NUMBER OF SEQ ID NOS: 10
```

```
; SOFTWARE: FaestSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589) ... (2583)
US-09-816-494-1

Alignment Scores:
Pred. No.: 2,18-194 Length: 3544
Score: 1552.00 Matches: 300
Percent Similarity: 99.67% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 4 Gaps: 0

US-10-029-345a-109_copy_1_302 (1-302) x US-09-816-494-1 (1-3544)
QY 1 MEALAHISGLUMETILEG1YTHRGINILEVALTHRG1ARGLEUVALIALEUWELUG1U 20
Db 589 ATGGCCCATGAGATGATGATGGAATCTCAATTTGTTACTGAGAGGTTGGCTGCTGGA 648
QY 21 SERGLYTHRGLULYSVALLEUWELI1EASPSERARGPROPHENVALG1UTYRANTHSER 40
Db 649 AGTGAACGAAAAAGGCTGCTTAATTTGATGATGACCGGCAATTTGGATACATACATCC 708
QY 41 H1S1LEUAGLUALIIEASNIEASNCYSERLYSEUWELYSARGALEUINGLN 60
Db 709 CACATTTGGAAAGCCATTAATATCACTGCTCCAAAGCTTTAGAGCAAGGTTGCAACG 768
QY 61 ASPLYVALLEUILETHRGULEUILEG1NHISERIALYSHIELYVALASP11LEAP 80
Db 769 GACAAAGTGTTAATTAACAAGCTTCATCAGCATTCAGCAACATTAAGTTGACATTTGAT 828
QY 81 CYSERGLNLYSVALVALTYRASPGLNSERSEGLNLYSVALIASERLEUSER 100
Db 829 TGCAGTCAGAAAGTTGATTAGATTACATCAAGCTCCAAAGATGCTCTCTCTCTTCA 888
QY 101 ASPCYSPHELEUHRVALLEUWELGLYLYSEUGLULYSERPHENASERVALHISLEU 120
Db 889 GACTGTTTCTCATCTGACTTCTGCGTMAACGAGAAAGCTTCACTCTGTTCACTG 948
QY 121 LEUVALAGLYG1PHEALAGUPHESERARGYSPHPROGLYLYSEU 140
Db 949 CTTCAGAGGTTGTTGCTGAGTTCTCTGTTGTTTCCCTGCGCTCTGTGAAGAAATCC 1008
QY 141 THRLEUVALPROTHRYSIIESERGLNPROCYSEUPROVALIAASNIILEG1YPROTHR 160
Db 1009 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGCTTACCTGTTGCCCAATTTGGCCAAAC 1068
QY 161 ARG1LEUAPROASNLEUTYRLEUGLYCYSGINARGASPVALLLEUASGLYLEUILE 180
Db 1069 CGAATCTTCCCAACTTATCTTGCGCTGCAGAGAGATGCTCCAAAGAGACTGATG 1128
QY 181 GING1NASGLY1LEG1YTYRVALLEUASNAIASERYTHRCPROLYSPROASPPHE 200
Db 1129 CAGCAGATGAGATGTTGATTATGTTAATGCCAGCAATACCTGTCCAAAGCTTGAATTTG 1188
QY 201 ILEPROGLNSERHISPHLEUARGVALPROVALIASNASERPHCYSG1ULYS1LEU 220
Db 1189 ATCCCGAATCTCATTTCTCGCTGCTGCTGATGATACAGCTTTTGTGAGAAATTTTG 1248
QY 221 PROTRPLEUASPLYSESERVALASPHLEI1EG1ULYSALYSALASERANGLYCYSPVA 240
Db 1249 CCGTGTTGGAACAATCAGTATGATTATTAAGAAAGCAAAAGCTCCCAATGATGATGTT 1308
QY 241 LEUVALHISCYSEUVALAG1Y1LESEARGSERIALATHRIEALIEALTYRI1EWET 260
Db 1309 CTAGTGCACTGTTTATGCTGGGATCTCCGCTCCGACCATTCCTATTCCTACATCATG 1368
QY 261 LYSATGMEIASPMETSERLEUASPG1UALATYRARGPHEVALYSG1ULYVARGPROTHR 280
```

```
Db 1369 AAGAGATGACATGCTTTAGATGAGCTTACGATTTGTGAAAGAAAAGACCTACT 1428
Qy 281 IISerProAnPheAnPheLeuGlyGlnLeuAspTyrGluYslyValIleValAn 300
Db 1429 ATATCTCAAACTTCAATTTCTGGGCAACTCTCGATGATGAGAAGATTAAAGAAC 1488
Qy 301 GlnThr 302
Db 1489 CAGACT 1494

RESULT 3
US-09-920-668-3
; Sequence 3, Application US/09920668
; Patent No. 6482644
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowserc
; APPLICANT: Bretc P. Monla
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE B EXPRESSION
; FILE REFERENCE: RTS-0246
; CURRENT APPLICATION NUMBER: US/09/920,668
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)...(2012)
US-09-920-668-3

Alignment Scores:
Pred. No.: 4,67e-120 Length: 2377
Score: 988.00 Matches: 190
Percent Similarity: 81.40% Conservative: 55
Best Local Similarity: 63.12% Mismatches: 54
Query Match: 63.17% Indels: 2
DB: Gaps: 2

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-920-668-3 (1-2377)
Qy 1 MetAlaHisGluMetIleGlyThrGlnIleVal---ThrGluArgValAlaIleLeu 19
Db 135 ATGGCTGGGAGCGGCTCCGAGAGAGTATGATGCCAAGAGCTGGCGACCTGCTG 194
Qy 20 GluSerGlyThrGluYsValLeuLeuIleAspSerArgProPheValGluYrAsnThr 39
Db 195 CGGGGCGGCGCTGGGGGCGGCTGGTGCATGACGCCGCTCTTCGTGGAGTACACAGC 254
Qy 40 SerHisIleLeuGluAlaIleAsnIleAsnCySerIleuMetIleYsValGArgLeuGln 59
Db 255 TGGCATGTGCTGAGCTCCGTCACATCTGCTGCCAAGCTGGAGGAGCGGCTGGCAG 314
Qy 60 GlnAspIleValLeuIleThrGluLeuIleGlnHisSerIleValIleValIle 79
Db 315 CAGGGCAGAGGTGACCATTTGGGAGCTCATCAGCCGCTCAGCAGCGGCTGGAGGCT 374
Qy 80 AspCySerGlnIleYsValIleValIleYrAspGlnSerSerGlnAspValIleSerLeuSer 99
Db 375 ACGGAGCCACAGACGCTGGTGTCTATGACCAAGACGCGGAGCGCCAGCTGTGGCC 434
Qy 100 SerAspCyPheLeuThrValLeuLeuGlyIleLeuGluYsSerPheAsnSerValHis 119
Db 435 GCGAGACGCTTCTCTCATCTGCTGAGCAGAGCTGAGCGGCTTCCAGACGCTGGCC 494
Qy 120 LeuLeuAlaIleYsPheAlaGluPheSerArgCyPheProGlyLeuCyGluGluYs 139
Db 495 ATCTCACTGGGGCTTGGCACCCTTCTCTCTGCTTCCCGGCTTCCGAGGGGAG 554
Qy 140 ---SerThrIleValProThrCyIleSerGlnProCyAspProValAlaAsnIleGly 158
Db 555 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 614
```

```
Qy 159 ProThrArgIleLeuProAnPheAnPheLeuGlyCyGlnArgAspValIleAsnIleVal 178
Db 615 CTGACCCGATCTGCTGCTCACTTACCTGGGCTCGCAGAGAGCTCTTAACAGAT 674
Qy 179 LeuIleGlnIleAsnGlyIleGlyYrValIleAsnAlaSerYrThrCyProIlePro 198
Db 675 CTGATGACCGAAATGGAATTAAGCTACGCTCTCAAGCCAGCACTCTGCCCCAAGCT 734
Qy 199 AspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCyGluYs 218
Db 735 GACTTCATCTGCGAGAGCGGCTTCATGCGGCTCCCATCAACACAACTACTGTGAAA 794
Qy 219 IleLeuProThrLeuAspIleYsSerValAspPheIleGlnIleValIleValIle 238
Db 795 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
Qy 239 CyValIleValHisCySerValIleGlyIleSerArgSerAlaThrIleAlaIleVal 258
Db 855 CAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 914
Qy 259 IleMetIleArgMetAspMetSerIleuAspGluAlaIleYrAspPheValIleYsVal 278
Db 915 ATCATGAGACCATGCGCATGCTCTCCGACAGCCTTACAGGTTGCTGAAGACAGGCGC 974
Qy 279 ProThrIleSerProAnPheAnPheLeuGlyGlnLeuAspTyrGluYsIle 298
Db 975 CCGTCCATCTCGGCCAACTTCACTTCTGGGCGACCTGCTGGATGACGACGCTG 1034
Qy 299 Lys 299
Db 1035 AAG 1037

RESULT 4
US-09-557-921-1
; Sequence 1, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Wei, Bo
; APPLICANT: Lucie, Ralf M.
; TITLE OF INVENTION: DSP-16
; FILE REFERENCE: 200125,416
; CURRENT APPLICATION NUMBER: US/09/557,921
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-557-921-1

Alignment Scores:
Pred. No.: 1,45e-51 Length: 1830
Score: 469.00 Matches: 107
Percent Similarity: 57.72% Conservative: 65
Best Local Similarity: 35.91% Mismatches: 96
Query Match: 29.99% Indels: 30
DB: Gaps: 7

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-557-921-1 (1-1830)
Qy 27 LeuLeuIleAspSerArgProPheValGluYrAsnThrSerHisIleLeuGluAlaIle 46
Db 603 GTCATCATTTAGTGAAGCCCTTCATGAGTACAAAGATGATCAATCCAAAGAGCTGTC 662
Qy 47 AsnIleAsnCySer---LysIleuMetIleYsValGArgLeuGlnIleAspIleValIle 65
Db 663 CACATTAACTGTGGCATTAAGATCAACCGCGAGACTCAAGAGGCGAAGATCACTGCT 722
Qy 66 ThrGluLeuIleGln---HisSerIleValIleValIleAspIleAspCySerGlnIle 84
Db 723 CTAGACTTATTTCTGTAGGAGGAGGAGACTCTTTCAAGAGATCTTTTCCAAAGAA 782
```

Qy 85 ValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSerAspCysPheLeu 104
Db 783 ATATAGTTATGATGAAATACCAATGAAACCAAGCGAGTGCCTCCACCACTT 842
Qy 105 ThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeuLeuAlaGly 124
Db 843 CACATAGCTCTGAGCTCCCTGAAAGAGAAAGCAAGAACTCTGGTGTGAAGGTGA 902
Qy 125 PheAlaGlnPheSerArgCysPheProGlyLeuGlyLys----- 139
Db 903 CTTAGTAGTTTAAACAGAACCAATGAAAACTCTGTGCAACTCCCTCCAGCTCCAAAG 962
Qy 140 -----SerThrLeuValProThrCysIleSer 148
Db 963 TGCCGGAGAGTGGGGGGGGCGGCATCCGGCCCTGAGCTTGTCACTCACTCACTC 1019
Qy 149 GlnProCysLeuPro--ValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyr 167
Db 1020 ---CCACACCACTGACATGAGAAAGCGTGAAGTCAACCCCATCTTGCCTCTGTC 1076
Qy 168 LeuGlyCysGlnAlaAspValLeuAsnLysGlnLeuIleGlnAsnGlyIleGlyTyr 187
Db 1077 CTGAGCAATGACGAGATGCTCAGACCTGAGACCAATGACGCGCTGAACATGAGCTAC 1136
Qy 188 ValLeuAsnAlaSer-----TyrThrCysProLysProAspPheIlePro 202
Db 1137 GTATCAACAGTCAACATCTTCCCTCTACCACTATGAAAGAGCTTTC----- 1190
Qy 203 GluSerHisPheLeuArgValProValAsnAspSerPheCysGlyLysIleLeuProTyr 222
Db 1191 ---AACTCAAGCCGCTGCCAGCCACTGACGAAACCAAGCAACTGCGCAGTAC 1244
Qy 223 LeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsnGlyCysValLeuVal 242
Db 1245 TTTAAAGAGGCTTTTGAATTGATGAGAGAGCTCAACAGTGGAAAGGGCTTCTCATC 1304
Qy 243 HisCysLeuAlaGlyLysSerArgSerAlaThrIleAlaIleAlaTyrIleMetLysArg 262
Db 1305 CACGCCAGAGCTGGGTCTCCCGCTCCGCCACCACTGCACTGCTTACTGATGAAGCAC 1364
Qy 263 MetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlyLysArgProThrIleSer 282
Db 1365 ACTGGATGACCATGACTGATGTTAATTGTCAAAGGCAAAAGCAACCAATATATCTCC 1424
Qy 283 ProAsnPheAsnPheLeuGlyGlnLeuAspTyrGlyLysIleLysAsn 300
Db 1425 CCAAACCTTAACCTTGAAGGCGAGTGTCTAGAGTTCAAGAAAGCACTAAACAC 1478

RESULT 5
US-09-016-434-1135
Sequence 1135, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434

FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1135:
SEQUENCE CHARACTERISTICS:
LENGTH: 2109 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1418933
US-09-016-434-1135

Alignment Scores:
Pred. No.: 2,91e-51 Length: 2109
Score: 467.50 Matches: 110
Percent Similarity: 53.56% Conservative: 63
Best Local Similarity: 34.06% Mismatches: 105
Query Match: 29.89% Indels: 45
Gaps: 7

US-10-029-345a-109_copy_1_302 (1-302) x US-09-016-434-1135 (1-2109)

Qy 19 LeuGluSerGlyThrGlyLysValLeuIleAspSerArgProPheValGlyTyrAsn 38
Db 430 CTGAGCTGGGCAACAGACGCGCTGCTGATGACTGCGCGGACGAGATACGAG 469
Qy 39 ThrSerHisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgTyrLeu 58
Db 490 TCGTGCACATCCAGTCGCGCATCAACGTGGCCATCCGGGATCATGTCGCGCGCTG 549
Qy 59 GlnGluAspLysValLeuIleThrGlnLeuIleGlnHisSerAla--LysHisLysVal 77
Db 550 CAGAAAGGTAACTGCGCGGTGCGCGCTCTTCAAGCGCGGAGACCGGACCTGTC 609
Qy 78 AspIleAspCys--SerGlnLysValValTyrAspGlnSerSerGlnAsp--Val 95
Db 610 ACCGGGCGCTGGGACCGACACAGTGTGTCTTACGACGAGACGACGACCTGAC 669
Qy 96 AlaSerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPhe 115
Db 670 GAGAAATACGGGCGGCGGAGTGTCTCGGCTGTCTCAAGAAAGCTCAAGACGAGGCG 729
Qy 116 AsnSerValHisLeuLeuAlaGlyLysPheAlaGlnPheSerArgCysPheProGlyLeu 135
Db 730 TGCCGGCGCTTACTGCTGAAAGGTGCTTCAAGATTCACACCGAGTCTCCCTGCAT 789
Qy 136 CysGlnGlyLysSerThrLeuValProThrCys--IleSerGlnProCysLeuProVal 154
Db 790 TGGAG--ACCATCTAAGCGGCTGTGTAGACGACGTCGCGCGCTTGGCAAGT 843
Qy 155 AlaAsnIleGly----- 158
Db 844 CTGGGGCTGGGGGCTGGGATCAGCTGATCTTCTCGGACATCGAGTCTGACCTT 903
Qy 158 ----- 158
Db 904 GACGAGACCCCAATAGTGCAACAGCTCGATGATGCTCCGCTGCCAACAGCAGCTT 963
Qy 159 ---ProThrArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsn 176
Db 964 TCCTTCCAGTGAAGATCTTCCCTTCTTACTTGGGCTGTGCAAAAGCTCCACCAAC 1023


```

QY 177 LysGluLeuIleGlnGlnAenglyIleGlyTyrValLeuAsnAlaSerThrCysPro 196
DB 1024 TTGGACGTGTGGAGCAATTCGACATCAATGACATTCAGCCCAATTTGCCG 1083
QY 197 LysPro---AspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPhe 215
DB 1084 ATCTCTTTGAGAACCGACGAGAGTTTAATACAGCAATCCCATCTCGATCCTGG 1143
QY 216 CysGluLysIleLeuProThrLeuAspLysSerValAspPheIleGluValAlaVal 235
DB 1144 AGCCAAACCTGTCACGATTTTCCCTGAGGCGCATTTCTTCATATGATGAAGCCGCGGC 1203
QY 236 SerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAla 255
DB 1204 AAGAACTGTGCTGTCTTGGTACATTCCTGGCTGGCAATGACCGCTCAGCTGACT 1263
QY 256 IleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLys 275
DB 1264 GTGGCTTACCTTATTCGAAAGCTCAATCTGTGATGACATGACATTCATTCGAA 1323
QY 276 GluLysArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlu 295
DB 1324 ATGAAATAATCCAAATATCCCTTAATCTCACTTCATGAGTCAAGTGTGAGCTTGA 1383
QY 296 LysLysIle 298
DB 1384 AGGACGCTG 1392

RESULT 6
US-09-023-655-946
Sequence 946, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Sellhammer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 946:
SEQUENCE CHARACTERISTICS:
LENGTH: 2109 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```

```

IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G1418933
US-09-023-655-946

Alignment Scores:
Score: 2,91e-51 Length: 2109
Percent Similarity: 467.50 Matches: 110
Best Local Similarity: 53.568 Conservative: 63
Query Match: 34.064 Mismatches: 105
DB: 29.894 Indels: 45
Gaps: 7

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-023-655-946 (1-2109)
QY 19 LeuGluSerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsn 38
DB 430 CTGAGACTGGGCAACGAGGCGCTGCTGATGAGACCTGCCGCGCAGAGCTATACGAG 489
QY 39 ThrSerHisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeu 58
DB 490 TGTCTGCACATCAGTGGCCATCAACGTGGCCATCCGAGCATCATGCTGCGGCGCTG 549
QY 59 GlnGlnAspLysValLeuIleThrGluLeuIleGlnHisSerAla---LysHisLysVal 77
DB 550 CAGAAAGGTAACTGCCGCGCTGCCGCGCTCTTCACGCGCGGAGAGACCGGACCCCTTC 609
QY 78 AspIleAspCys---SerGlnLysValValTyrAspGlnSerSerGlnAsp---Val 95
DB 610 ACCCGCGCTGTGGACACGACACAGTGTCTTACGACGAGAGAGAGAGAGAGAGAGAG 669
QY 96 AlaSerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPhe 115
DB 670 GAGAAATCAGGCGCGCGCGAGTCTGTGCGGCGCTGCTCAAGAAAGTCAAGAGAGAGG 729
QY 116 AsnSerValHisLeuLeuValGlyGlyPheAlaGluPheSerArgCysPheProGlyLeu 135
DB 730 TGGCGGCGCTTCTACTGGAAGTGGCTTCAGTAATTCAGACCGAGTCTCCCTGCAT 789
QY 136 CysGluLysLysSerThrLeuValProThrCys---IleSerGlnProCysLeuProVal 154
DB 790 TGGAG---ACCAATCTAAGACGCGCTGTGAGAGAGAGCTCGCGCGCTGCGAGT 843
QY 155 AlaAsnIleGly----- 158
DB 844 CTGGGCGCTGGGCGCTGGGAGATCAAGCTCTGACTCTTCCTCGGACATCGAGTGAAGCTT 903
QY 158 ----- 158
DB 904 GACCGAGACCCCAATGTCACACAGCTCGGATGATGTCCTGCTGCCAAGCCAGCCCT 963
QY 159 ---ProThrArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsn 176
DB 964 TCTTCTCCAGTGGAGATCTTGCCCTTCTCTACTTGGCTGTGCCAAAGCTCCACCAAC 1023
QY 177 LysGluLeuIleGlnGlnAenglyIleGlyTyrValLeuAsnAlaSerThrCysPro 196
DB 1024 TTGGACGTGTGGAGCAATTCGACATCAATGACATTCAGCCCAATTTGCCG 1083
QY 197 LysPro---AspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPhe 215
DB 1084 ATCTCTTTGAGAACCGACGAGAGTTTAATACAGCAATCCCATCTCGATCCTGG 1143
QY 216 CysGluLysIleLeuProThrLeuAspLysSerValAspPheIleGluValAlaVal 235
DB 1144 AGCCAAACCTGTCACGATTTTCCCTGAGGCGCATTTCTTCATATGATGAAGCCGCGGC 1203
QY 236 SerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAla 255
DB 1204 AAGAACTGTGCTGTCTTGGTACATTCCTGGCTGGCAATGACCGCTCAGCTGACT 1263
QY 256 IleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLys 275

```

```

Db      1264 GGGCTTACTTATGACGAGCTCAATCTGTGATGACAGAGCTTATGACATTGTCAAA 1323
Oy      276 GtUlyeArXpProThrIleSerProaenpheanpheleuGlyGlnleuAaPtyrglu 295
Db      1324 ATGAAAAAATCCAAACATATCCCTACTTCAACTTCATGCGTGCAGCTGTGACATTGCAG 1383
Oy      296 LylLyAile 298
Db      1384 AGGAGCGCTG 1392

RESULT 7
US-09-016-434-1100
; Sequence 1100, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 845-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G1255784
; US-09-016-434-1100

Alignment Scores:
Pred. No.: 1e-45 Length: 2240
Score: 426.00 Matches: 100
Percent Similarity: 54.85% Conservative: 64
Best Local Similarity: 33.44% Mismatches: 113
Query Match: 27.24% Indels: 22
DB: 4 Gaps: 9

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-016-434-1100 (1-2240)
Oy      19 LeuGisergYtThnGtUlySValleuLeuIleAaSpSerXpProPheValGtUlyrAa 38
Db      234 CTGCCGAGCGCG---GGCAAGTGGCTGTGCTGAGACTGCAAGACCGCTTCTGGCGCACAGC 290
Oy      39 ThrsrHsIstleuGtUlaIleAsnIleAsnCySerIySleuWctlySArGArGleu 58

```

[illegible]

```

Db      870  CAGAGCGGAGTATCATCTCCCGGAATTCCAGTTTCATGGGCCAGCTGCTGACGATTGTAG 929
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Cy      296  Lysbyslie 298
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      930  TCCCAAGTG 938
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

RESULT 9
US-08-530-290-11
/ Sequence 11, Application US/08530290
/ Patent No. 5958721
/ GENERAL INFORMATION:
/ APPLICANT: Marshall, Christopher John
/ APPLICANT: Ashworth, Alan
/ APPLICANT: Hughes, David Anthony
/ TITLE OF INVENTION: Methods for Screening of Substances for
/ TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
/ NUMBER OF SEQUENCES: 24
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/530,290
/ FILING DATE: 14-DEC-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO PCT/GB94/00694
/ FILING DATE: 31-MAR-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9402573.1
/ FILING DATE: 10-FEB-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9307250.2
/ FILING DATE: 07-APR-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Baerlan, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 084611-000000US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1238 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/ US-08-530-290-11

Alignment Scores:
Pred. No.: 5,026-46 Length: 1238
Score: 425.00 Matches: 99
Percent Similarity: 53.85% Conservative: 62
Best Local Similarity: 33.11% Mismatches: 116
Query Match: 27.17% Indels: 22
DB: 2 Gaps: 8

US-10-029-3454-109_COPY_1_302 (1-302) * US-08-530-290-11 (1-1238)
Cy      19  leuglusercIyThrcIuYsValIeuleuIleApsSerAqPpoeValgIuTYrAzn 38
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      168  CTGCGGAGCGCGC---CGCAAGTGCTGTGCTGCGACGTCAGACCGGTTCTTGCGCGCAGC 224
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

39 ThrSerHisIleLeuGluAlaIleAsnIleAsnCySerIysLeuMetIysArgArgLeu 58
Db 225 GCGGGCTACATCCGATCGGTGCGTCAACGCGCTGTAAACACCATCGTGGCGCGG--- 281
Cy 59 GlnGlnAspIysValIleuIleThrGluLeuIle-----GlnHisSerAlaIysHisIys 76
Db 282 GCTAAGAGGCTCCGTGAGCTGTGAGACAGATCCTGCGCCGAGGAGGAGGTACGGCGCCG 341
Cy 77 ValAspIleAspCySerGlnIysValIValIValIValIValIValIValIValIValI 96
Db 342 TTGGCTCCCGGCTCTACTCGCGCGTCAATCGTCTACAGACGACCGCGCGCGCGAG 401
Cy SerLeuSerSerAspCySerPheLeuThrValLeuLeuGlyIysLeuGlyIysSerPhe--- 115
Db 402 AGCTCCCGGAGGACGACGCGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 461
Cy 116 ---AsnSerValHisIleuLeuAlaIleGlyIysPheAlaGluPheSerArgCysPheProGly 134
Db 462 CGCACCGCATCTCGCTGCTCAAGCGGCTATGAGAGGTTTCTCCGAGTACCCAGAA 521
Cy 135 LeuCySerGluGlyIysSerThrLeu-----ValProThrCysIleSer 148
Db 522 TTCTGTTCTAAACCAAGCGCTGTGACACCATCCACCGCGGTTCGCCCGACGACCA 581
Cy 149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
Db 582 GAGCCCTTGACCTGAGCTGAGCTCGTGTGGAGCCCATACAGACGACGAGGTCCT 641
Cy 160 ThrArgIleLeuProAsnLeuIleuGlyCysGlnArgAspValLeuAsnIysGluLeu 179
Db 642 GTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 701
Cy 180 IleGlnGlnAsnGlyIleGlyIysValIleuAsnIleSerIysThrCysProIysProAsp 199
Db 702 CTGAGCGCCCTGGGACATACCGCTGTGTGATGTCTCTCTCTCTCTCTCTCTCTCTCT 758
Cy 200 PheIleProGlnIysSerHisPheLeuArgValProValAsnAspSerPheCysGluIysIle 219
Db 759 TTGTAAGAGACATCACTAGTCAATGATCCCACTGTGAAGATTAACCAAGCGGACATC 818
Cy 220 LeuProIleLeuAspIysSerValAspPheIleGlyIysAlaIysAlaIysSerAsnGlyCys 239
Db 819 AGCTCCGTGTTCATGAGGACCATAGATCACTCATGCGGTGAGAGGACTGCGGTGGCGC 878
Cy 240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIysIle 259
Db 879 GTGCTGTGACATCCGACGCGGACATCTCGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 938
Cy 260 MetIysArgMetAspMetSerLeuAspGluAlaIysArgPheValIysGluIysArgPro 279
Db 939 ATGATGAAGAAACGGGTGAGGCTGAGAGGCTTCGAGTTGTTAAGCAGCGCGCAGC 998
Cy 280 ThrIleSerProAsnPheAsnPheLeuGlyIleuLeuLeuAspIysArgIysIysIle 298
Db 999 ATCATCTCGCCCACTTACGCTCATGCGGAGCTGTGAGTTCGAGTCCAGGTG 1055

RESULT 10
US-09-702-705-803
Sequence 803 Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedrick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14

CURRENT APPLICATION NUMBER: US/09/702, 705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 803
LENGTH: 1238
TYPE: DNA
ORGANISM: Homo sapien
US-09-702-705-803

Alignment Scores:
Pred. No.: 5, 02e-46 Length: 1238
Score: 425.00 Matches: 99
Percent Similarity: 53.85% Conservative: 62
Best Local Similarity: 33.11% Mismatches: 116
Query Match: 27.17% Indels: 22
Gaps: 8

US-10-029-345a-109_copy_1_302 (1-302) x US-09-702-705-803 (1-1238)

19 LeuGluSerGlyThrGluIysValIleuLeuIleAspSerArgPheValIleGlyIysAsn 38
Db 168 CTGCGAGCGGCG---GCGATGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 224
Cy 39 ThrSerHisIleLeuGluAlaIleAsnIleAsnCySerIysLeuMetIysArgArgLeu 58
Db 225 GCGGGCTACATCCGATCGGTGCGTCAACGCGCTGTAAACACCATCGTGGCGCGG--- 281
Cy 59 GlnGlnAspIysValIleuIleThrGluLeuIle-----GlnHisSerAlaIysHisIys 76
Db 282 GCTAAGAGGCTCCGTGAGCTGTGAGACAGATCCTCCGCGGAGGAGAGGTACCGCGCGC 341
Cy 77 ValAspIleAspCySerGlnIysValIValIValIValIValIValIValIValIValI 96
Db 342 TTGGCTCCCGGCTCTACTCGCGCGTCAATCGTCTACAGACGACCGCGCGCGAG 401
Cy 160 ThrArgIleLeuProAsnLeuIleuGlyCysGlnArgAspValLeuAsnIysGluLeu 179
Db 642 GTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 701
Cy 180 IleGlnGlnAsnGlyIleGlyIysValIleuAsnIleSerIysThrCysProIysProAsp 199
Db 702 CTGAGCGCCCTGGGACATACCGCTGTGTGATGTCTCTCTCTCTCTCTCTCTCTCTCT 758
Cy 200 PheIleProGlnIysSerHisPheLeuArgValProValAsnAspSerPheCysGluIysIle 219
Db 759 TTGTAAGAGACATCACTAGTCAATGATCCCACTGTGAAGATTAACCAAGCGGACATC 818
Cy 220 LeuProIleLeuAspIysSerValAspPheIleGlyIysAlaIysAlaIysSerAsnGlyCys 239
Db 582 GAGCCCTTGACCTGAGCTGAGCTCGTGTGGAGCCCATACAGACGACGAGGTCCT 641
Cy 160 ThrArgIleLeuProAsnLeuIleuGlyCysGlnArgAspValIleuAsnIysGluLeu 179
Db 642 GTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 701
Cy 180 IleGlnGlnAsnGlyIleGlyIysValIleuAsnIleSerIysThrCysProIysProAsp 199
Db 702 CTGAGCGCCCTGGGACATACCGCTGTGTGATGTCTCTCTCTCTCTCTCTCTCTCTCT 758
Cy 200 PheIleProGlnIysSerHisPheLeuArgValProValAsnAspSerPheCysGluIysIle 219
Db 759 TTGTAAGAGACATCACTAGTCAATGATCCCACTGTGAAGATTAACCAAGCGGACATC 818
Cy 220 LeuProIleLeuAspIysSerValAspPheIleGlyIysAlaIysAlaIysSerAsnGlyCys 239
Db 819 AGCTCCGTGTTCATGAGGACCATAGATCACTCATGCGGTGAGAGGACTGCGGTGGCGC 878
Cy 240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIysIle 259
Db 879 GTGCTGTGACATCCGACGCGGACATCTCGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 938
Cy 260 MetIysArgMetAspMetSerLeuAspGluAlaIysArgPheValIysGluIysArgPro 279

```
DB 939 ATGATAGAGAAACGGGAGGCTGTGAGAGAGCCCTTGAGTTCCTTAAGACGCGCCAGC 998
QY 280 ThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnIlyblyle 298
DB 999 ATCATCTCGCCCAACTTCACTTCATGCGGAGACCTGCTGCAAGTTCCAGTCCAGGTG 1055

RESULT 11
US-09-736-457-803
/ Sequence 803, Application US/09736457
/ Patent No. 6509448
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ APPLICANT: Wang, Aijun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C15
/ CURRENT APPLICATION NUMBER: US/09/736,457
/ CURRENT FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 1864
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 803
/ LENGTH: 1238
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-09-736-457-803

Alignment Scores:
Pred. No.: 5,02e-46 Length: 1238
Score: 425.00 Matches: 99
Percent Similarity: 53.85% Conservative: 62
Best Local Similarity: 33.11% Mismatches: 116
Query Match: 27.17% Indels: 22
Gaps: 8
DB: 4

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-736-457-803 (1-1238)
QY 19 LeuGlnSerGlyThrGlnIlyValLeuLeuIleAspSerArgProPheValGlnTyrAsn 38
DB 168 CTCGCCAGCGGC---GGCAAGTGCCTGCTGCTGAGACTGACACCGTTCTCGGCGCACAGC 224
QY 39 ThrSerHisIleLeuGlnIleAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 58
DB 225 GCGGCGCTACATCTTGAAGTGTGTCACGTCGCTGTAAACCAATCGCGCGCGCGG--- 281
QY 59 GlnGlnAspIlyValLeuIleThrGlnIleIle-----GlnHisSerAlaIlyblyle 76
DB 282 GCTAAGGCTCCCTGAGCTGAGACAGATCTGCGCCGCGACAGAGAGTACCGCCCGC 341
QY 77 ValAspIleAspCysSerGlnIlyValValIlyTyrAspGlnSerSerGlnAspValAla 96
DB 342 TTGCGCTCCGCGCTTACTCGCGCGCTCATGCTTACAGACGAGCGGCGCGCGCGAG 401
QY 97 SerLeuSerSerArgPheLeuThrValLeuLeuGlnIlyblyleGlnIlyblySerPhe--- 115
DB 402 AGCCTTCGCGAGACAGACCGTGTGCTGTGTGAGCGGCTGCGCGCGCGCGAG 461
QY 116 ---AsnSerValHisLeuLeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGly 134
DB 462 CGCAGCGACATCTGCGCTGTCAAGGCGGTATGAGAGTTTCTCCGAGTACCGAGAA 521
QY 135 LeuGlnGlnIlyblySerThrLeu-----ValProThrCysIleSer 148
DB 522 TTCTGTTCTTAAACCAAGGCGCTGCGAGCATCCACCCCGGTTCCCGCCAGCGCACCA 581
```

```
QY 149 GlnPro-----CysLeuProValAla-----AsnIleGlyPro 159
DB 582 GAGCCCTTGAGCTGAGCTGAGCTCTCTGTGGAGACCCCACTACAGACAGAGAGGTCT 641
QY 160 ThrArgIleLeuProAsnLeuTyrIleuGlyCysGlnIlyAspValLeuAsnIlyblyleu 179
DB 642 GTGAGATCTCTTCCCTTCTTACTGCTGAGAGTCCCTTACCATGCTGCGCGGAGAGCATG 701
QY 180 IleGlnIleAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProIlyProAsp 199
DB 702 CTGAGCGCCCTGGGCACTACGCGCTGTGTAAATGTCCTCGGACTGCCAACC---CAC 758
QY 200 PheIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIlyblyle 219
DB 759 TTGGAAGACACTATATAGTACAAAGTCCATCCAGTGAAGATACCAACAGCGCGCATC 818
QY 220 LeuProTyrLeuAspIlySerValAspPheIleGlnIlyAlaIlyblyleAsnArgIlyCys 239
DB 819 AGCTCTGCTTCAATGAGACCATAGATCATGATGATGATGATGATGATGATGATGATGATG 878
QY 240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 259
DB 879 GTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 938
QY 260 MetIlyArgMetAspMetSerLeuAspGlnIlyAlaTyrArgPheValIlyblyleArgPro 279
DB 939 ATGATAGAGAAACGGGTGAGGCTGTGAGAGGCTTGAATTCCTTAAGACGCGCCAGC 998
QY 280 ThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnIlyblyle 298
DB 999 ATCATCTCGCCCAACTTCACTTCATGCGGAGACCTGCTGCAAGTTCCAGTCCAGGTG 1055

RESULT 12
US-09-614-124B-803
/ Sequence 803, Application US/09614124B
/ Patent No. 6630574
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ APPLICANT: Wang, Aijun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ FILE REFERENCE: 210121.478C9
/ CURRENT APPLICATION NUMBER: US/09/614,124B
/ CURRENT FILING DATE: 2001-07-11
/ NUMBER OF SEQ ID NOS: 1668
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 803
/ LENGTH: 1238
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-09-614-124B-803

Alignment Scores:
Pred. No.: 5,02e-46 Length: 1238
Score: 425.00 Matches: 99
Percent Similarity: 53.85% Conservative: 62
Best Local Similarity: 33.11% Mismatches: 116
Query Match: 27.17% Indels: 22
DB: 4

US-10-029-345A-109_COPY_1_302 (1-302) x US-09-614-124B-803 (1-1238)
QY 19 LeuGlnSerGlyThrGlnIlyValLeuLeuIleAspSerArgProPheValGlnTyrAsn 38
DB 168 CTCGCCAGCGGC---GGCAAGTGCCTGCTGCTGAGACTGACACCGTTCTCGGCGCACAGC 224
QY 39 ThrSerHisIleLeuGlnIleAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 58
```



```
DB 939 ATGATGAAGAAACGGGTGAGGCTGAGAGCCCTTCAGTTCCGTTAAGACAGCCGCGAGC 998
QY 280 ThrIleSerProAnpPheAnpPheLeuGlyValLeuLeuAspTryArgIlyValIle 238
DB 999 ATCATCTCGCCCACTTCACTTCATGAGGCGAGCTGCTGAGCTTCGAGTCCAGAGTG 1055

RESULT 14
US-09-589-184-803
; Sequence 803, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaltanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 803
; LENGTH: 1238
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-589-184-803

Alignment Scores:
Pred. No.: 5.02e-46 Length: 1238
Score: 425.00 Matches: 99
Percent Similarity: 53.85% Conservative: 62
Best Local Similarity: 33.11% Mismatches: 116
Query Match: 27.17% Indels: 22
Gaps: 8

US-10-029-345a-109_copy_1_302 (1-302) x US-09-589-184-803 (1-1238)
QY 19 LeuGluSerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTryAsn 38
DB 168 CTCGCCAGCGCCG---GCCAGTCTCTCTCTCTGATCGACAGACCTTCTCGGCGACAGC 224
QY 39 ThrSerHisIleLeuGluAlaIleAsnIleAsnCySerIlyValMetIlyArgArgLeu 58
DB 225 GCGGGCTACATCTTCAAGTTCGCTCAACGTCGCTGTAACACATCGTGGCGCGCGG--- 281
QY 59 GluGlnAspIlyValLeuIleThrGluLeuIle-----GlnHisSerAlaIlyValIle 76
DB 282 GCTAAGGCTCTCGGAGCTCGAGCAGATCTCGCCGCGAGAGAGAGTACGCGCCGC 341
QY 77 ValAspIleAspCySerGlnIlyValIleValIlyAspGlnSerSerGlnAspValAla 96
DB 342 TTGCGCTCCGCGCTCTACTCGGCGGTATCTCTACACAGACGCGACCGCGCGCGAG 401
QY 97 SerLeuSerSerArgCyPheLeuThrValLeuLeuGlyIlyValLeuGlyIlySerPhe--- 115
DB 402 AGCCTCGCGAGAGACACACCGTGTGCTGTGTGTGAGCGCGCGCGCGCGAG 461
QY 116 ---AsnSerValHisLeuLeuAlaGlyIlyPheAlaGluPheSerArgCyPheProGly 134
DB 462 CGACACCGACATCTGCTCTCAAGCGGCTATGAGGTTTCTCCGAGTCCCGAGAA 521
QY 135 LeuGluSerGlyIlySerThrLeu-----ValProThrCyIleSer 148
DB 522 TTCTGTCTTAAACAGAGCCCTTGAGCGCATCCACCCCGGTTCCCGCCAGCGCAC 581
QY 149 GlnPro-----CyLeuProValAla-----AsnIleGlyPro 159
DB 582 GAGCCCTTGGACCTGAGACTGAGCTCTGTGGAGCCCGCACTACAGACAGAGAGGTCCT 641
```

```
QY 160 ThrArgIleLeuProAnpLeuIlyrLeuGlyCyGlnArgAspValLeuAsnIlyGluLeu 179
DB 642 GTGAGATCTTCTCTCTCTCTCTCTGAGAGCTTACCATGCTGCGCGAGAGACATG 701
QY 180 IleGlnGlnAsnGlyIlyGlyrValLeuAsnAlaSerTryrThrCyProIlyProAsp 199
DB 702 CTGAGCGCCTGGGACATCAGCGCTGTGTAAATGTCCTCTCGAGCTCCCAAC---CAC 758
QY 200 PheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCyGluIlyIle 219
DB 759 TTGAAGACACTATCACTAGTCAAGTGCATCCAGTGAATATACCAAGCGCGACATC 818
QY 220 LeuProThrLeuAspIlySerValAspPheIleGluIlyValIlyValAsnArgIlyCys 239
DB 819 AGCTCTGTCTTACATGAGCATATGATCATGATCCGTGAAGACTCGCTGGCGCC 878
QY 240 ValLeuValHisCySerLeuAlaGlyIlySerArgSerAlaThrIleAlaIleAlaTryIle 259
DB 879 GTCTGTGTCATGCGCGAGCGGCGATCTCGCGGTGCGCCACATCTGCTGCGCTACTG 938
QY 260 MetIlyArgMetAspMetSerLeuAspGluAlaTryrArgPheValIlyGlyIlyArgPro 279
DB 939 ATGATGAAGAAACGGGTGAGGCTGAGAGGCTTCGAGTTCGTTAAGACAGCCGCGAGC 998
QY 280 ThrIleSerProAnpPheAnpPheLeuGlyValLeuLeuAspTryArgIlyValIle 238
DB 999 ATCATCTCGCCCACTTCACTTCATGAGGCGAGCTGCTGAGCTTCGAGTCCAGAGTG 1055

RESULT 15
US-09-702-705-825
; Sequence 825, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaltanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 825
; LENGTH: 2064
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-825

Alignment Scores:
Pred. No.: 1.18e-45 Length: 2064
Score: 425.00 Matches: 100
Percent Similarity: 54.85% Conservative: 64
Best Local Similarity: 33.44% Mismatches: 113
Query Match: 27.17% Indels: 22
Gaps: 9

US-10-029-345a-109_copy_1_302 (1-302) x US-09-702-705-825 (1-2064)
QY 19 LeuGluSerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTryAsn 38
DB 511 CTCGCCAGCGCCG---GCCAGTCTCTCTCTCTGATCGACAGACCTTCTCGGCGACAGC 567
QY 39 ThrSerHisIleLeuGluAlaIleAsnIleAsnCySerIlyValMetIlyArgArgLeu 58
DB 568 GCGGGCTACATCTTCAAGTTCGCTCAACGTCGCTGTAACACATCGTGGCGCGCGG--- 624
```


QY 59 GlnGlnAspIysValLeuIleThrGluLeuIle-----GlnHisSerAlaIysHisLys 76
 ::: ||| ::: ::: ::: :::
Db 625 GCTAAGGGCTCCTGTAGAGCTGAGAGATCTCCCGGAGAGAGAGTACGCGCCGAC 684
QY 77 ValAspIleAspCysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAla 96
 ::: ||| ||| ||| ||| |||
Db 685 TTGCGCTCCGCTTACTCTGCGGCTCATGCTTACAGAGAGCGAGCCGCGCGCCGAG 744
QY 97 SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPhe--- 115
 ||| ||| ||| ||| ||| |||
Db 745 AGCCTTCGCGAGGACAGCACCGGTGCTGTGTGTGTCAGGCGCTGTGCGCAACGCCGAG 804
QY 116 ---AsnSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGly 134
 ::: ||| ||| ||| ||| |||
Db 805 CGACCGACATCTGCTGCTCAAGCGGCTATGAGAGGTTTCTCCAGTACCCAGAA 864
QY 135 LeuCysGluGlyLysSerThrLeu-----ValProThrCysIleSer 148
 ||| ||| ||| ||| ||| |||
Db 865 TTCTGTCTTAAACCAAGGCCCTGGCAGCCATCCACCCCGGTTTCCCGCAGTGCACA 924
QY 149 GlnPro-----Cys---LeuProValAlaAsnIle---GlyPro 159
 ::: ||| ||| ||| ||| |||
Db 925 GAGCCCTTGACCTGGGCTGCACTCTCTGTGGGACCCCACTACAGACCGAGGCGGCTCT 984
QY 160 ThrArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 179
 ||| ||| ||| ||| ||| |||
Db 985 GTGAGATCTCTCCCTCTCTACCTCGGCACTGCTACCATCCATGCTGCCCGAGAGACATG 1044
QY 180 IleGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProLysProAsp 199
 ::: ||| ||| ||| ||| |||
Db 1045 CTGAGCGCCCTGGGATCACGGCTGTGTGAATGTCTCTCGACTGCGCCAAC---CAC 1101
QY 200 PheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle 219
 ||| ||| ||| ||| ||| |||
Db 1102 TTGGAAGCACATATCAGTACAGTCCAGTGTGAAGATACACAGGCGGACATC 1161
QY 220 LeuProTyrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCys 239
 ||| ||| ||| ||| ||| |||
Db 1162 AGCTCCTGTTCATGGAAGCATAGATCATCATGATCCCTGAAGACTGCGGTGGGCGC 1221
QY 240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 259
 ||| ||| ||| ||| ||| |||
Db 1222 GTGCTGTGTGCTGCTGCGGCGATCTGCGGCTGCGGCGCCATCTGCTGCTTACCTG 1281
QY 260 MetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgPro 279
 ||| ||| ||| ||| ||| |||
Db 1282 ATGATGAAGAAACGGGTGAGGCTGAGAGAGGCTTTCGAGTTCTTAAAGCAGCGCGCAGC 1341
QY 280 ThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLysIle 298
 ||| ||| ||| ||| ||| |||
Db 1342 ATCATCTCGCCCACTTACGCTTCAATGGGGGAGCTGTGCAATTTCAGATCCACAGGTG 1398

Search completed: June 22, 2004, 02:46:21
Job time : 66.9628 secs